INTRODUCCION AL ANALISIS MATEMATICO

ROBERT G. BARTLE
Profesor de Matemáticas
Universidad de Illinois
Urbana-Champaign



EDITORIAL LIMUSA
MEXICO 1980

Find your solutions manual here!

Versión autorizada en español de la
obra publicada en inglés por
John Wiley & Sons, Inc., bajo el título de
THE ELEMENTS OF REAL ANALYSIS 2a. Edition.
 ISBN 0 – 471 – 05464 – X

Versión española:
MA. CRISTINA GUTIERREZ GONZALEZ
Matemática y Profesora de Matemáticas
de la Facultad de Ciencias de la Universidad
Nacional Autónoma de México.

Revisión:
ALEJANDRO JAVIER DIAZ BARRIGA CASALES
Doctor en Matemáticas, Investigador de Tiempo
Completo y Catedrático de Matemáticas de la
Universidad Nacional Autónoma de México

Balderas 95, Primer piso, México 1, D.F.
Miembro de la Cámara Nacional de la
Industria Editorial. Registro Núm. 121

Primera edición: 1980
Impreso en México
(1937)

ISBN 968 – 18 – 0997 – 1

i005 (653x178x2 pcx)

A mis padres

Prólogo

Hasta hace unos años se podía esperar que un estudiante de matemáticas avanzadas desarrollara la técnica para resolver problemas que implicaban gran cantidad de cálculo; sin embargo, no era de esperar que dominara "sutilezas teóricas" como convergencia uniforme o continuidad uniforme. Se requería que pudiera emplear el teorema de la función implícita, pero no que conociera sus hipótesis. Esta situación ha cambiado; en la actualidad se considera importante que todos los estudiantes avanzados en matemáticas: futuros matemáticos, especialistas en computación, físicos, ingenieros o economistas, dominen la naturaleza teórica básica de la materia. De este modo entenderán la fuerza, así como la limitación de la teoría general.

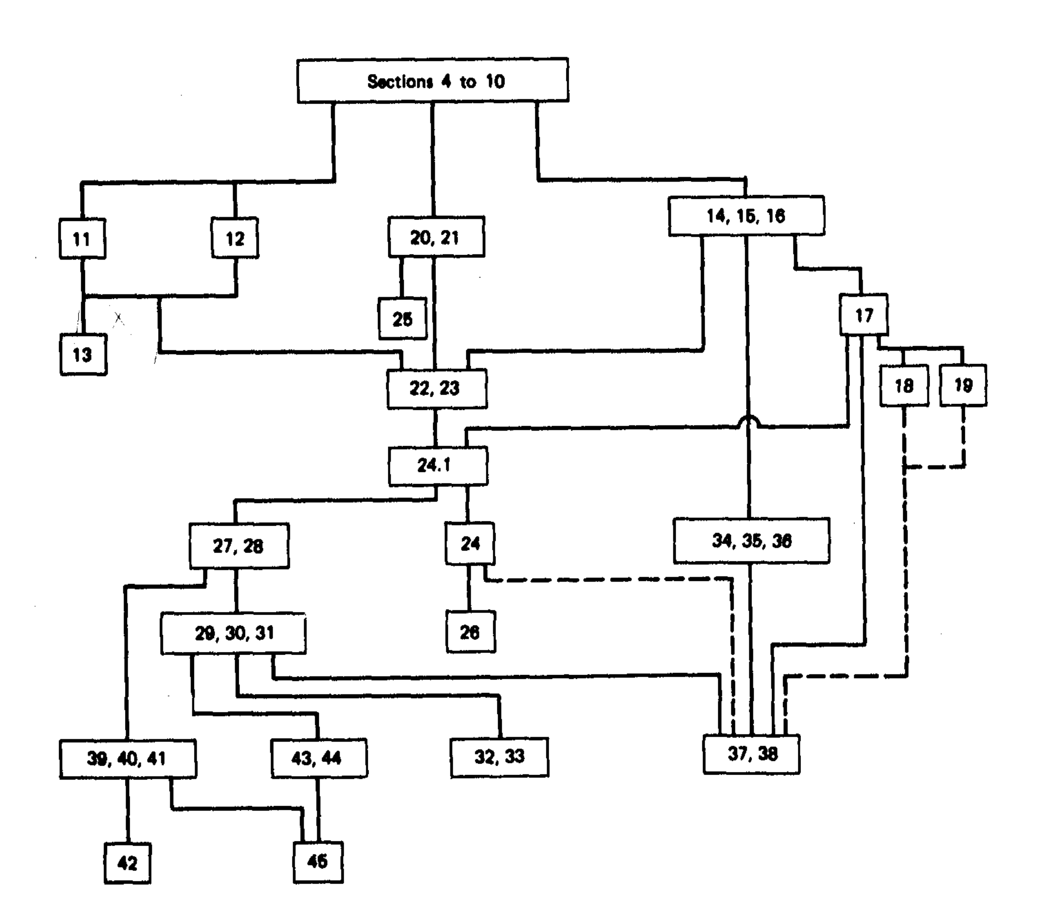
Este libro de texto se pudo realizar gracias a la experiencia obtenida al impartir clases de análisis matemático en la Universidad de Illinois desde 1955.

Mi público por lo general varía desde estudiantes universitarios del primer año, extraordinariamente bien preparados, hasta estudiantes graduados. Por lo general, la mayor parte de ellos no se especializa en matemáticas, pero ha estudiado cuando menos el equivalente a tres semestres de cálculo (no riguroso), incluyendo derivadas parciales, integrales múltiples, integrales de línea y series infinitas. Es conveniente para todos los estudiantes haber cursado un semestre de álgebra lineal o moderna para poder entender este curso en el que se demuestran teoremas analíticos. Sin embargo, puesto que muchos de los estudiantes no poseen tales bases, comienzo el estudio del análisis con unas cuantas demostraciones algebraicas para prepararlos.

En esta edición, en las secciones 4 y 5, establezco las propiedades algebraicas y de orden del sistema de números reales, de una manera más simplificada que en la primera edición. Además, en la sección 8 doy a conocer las definiciones de espacio vectorial y espacio normado, ya que estos conceptos aparecen con mucha frecuencia en las matemáticas modernas. Asimismo, he abreviado algunas secciones con el objeto de hacer más accesible el material y de proporcionar flexibilidad adicional al utilizar este libro como texto. He agregado muchos ejercicios y proyectos nuevos procurando mantener el libro al mismo nivel de complejidad que en la primera edición. En la primera parte del libro ha habido pocos

cambios. No obstante, dado que la experiencia ha demostrado que el análisis de diferenciación e integración en R^p fue demasiado breve en la primera edición, integré la teoría de funciones de una variable en un solo capítulo y extendí considerablemente el estudio de funciones de más de una variable.

En las secciones 1 a 3 doy la terminología y notación de teoría de conjuntos, que se emplean posteriormente, y establezco algunos conceptos básicos. Sin embargo, estas secciones *no* proporcionan una presentación sistemática de la teoría de conjuntos. (Dicha presentación no es necesaria ni deseable en esta etapa). Estas secciones se deberán revisar brevemente y consultar más adelante en caso necesario. En realidad, el texto empieza en la sección 4, y en la sección 6 se introduce "el análisis". En un semestre es posible abarcar las secciones 4 a 12, 14 a 17, 20 a 24.1 y la mayor parte de las secciones 27 a 31. Yo haría uso del criterio del profesor para impartir brevemente algunos otros temas (como series) a cambio de moderar (o incluso omitir) varios resultados que no son necesarios para el material subsecuente. Dado que el texto completo comprende un poco más del material que normalmente se podría abarcar en un año, a este nivel, es posible que el profesor restrinja la exposición de algunas secciones. Sin embargo, será útil para el estudiante disponer del material adicional como material de



consulta posterior. La mayoría de los temas generalmente asociados con cursos de "cálculo avanzado" se tratan aquí. La principal excepción es el tema sobre integrales de línea y de superficie y el teorema de Stokes; estos puntos no se tratan ya que un análisis intuitivo es propiamente parte del cálculo y un análisis riguroso requiere de un estudio bastante amplio para que resulte útil.

La dependencia lógica entre las distintas secciones de este libro de texto se explica por medio del diagrama adjunto. Una línea sólida en este diagrama indica dependencia en la sección anterior; una línea discontinua indica una ligera dependencia. Por ejemplo, todas las definiciones, teoremas, corolarios y lemas están numerados consecutivamente de acuerdo con el número de la sección. He asignado nombres a los teoremas más importantes, siempre que haya resultado apropiado algún nombre. Las demostraciones se destacan del texto por medio del encabezado DEMOSTRACION y la terminación Q.E.D.

No es necesario subrayar demasiado la importancia de los ejercicios y los proyectos; sólo empleando grandes esfuerzos concentrados en su resolución existirá la posibilidad de dominar el material de este libro. Los proyectos desarrollan un tema específico en una sucesión encadenada de ejercicios; pensamos que le transmiten al estudiante, cuando menos, una muestra del placer (¡el tormento!) de hacer investigación en matemáticas. Espero que ningún estudiante deje de hacer la prueba con varios de estos proyectos ya que, a mi parecer, es una fase particularmente valiosa de este libro.

Al escribir este libro he utilizado mi experiencia del salón de clases y he recibido influencia de muchas fuentes. Me he beneficiado discutiendo con estudiantes y colegas, y desde la publicación de la primera edición he mantenido bastante correspondencia con estudiantes y maestros de otras instituciones. Doy las gracias a todos aquellos que han hecho comentarios y sugerencias; su interés por mejorar el libro me ha estimulado para llevar a cabo esta revisión. Los profesores K.W. Anderson, W.G. Bade y A. L. Peressini leyeron el manuscrito de la primera edición y aportaron útiles sugerencias. Agradezco especialmente a mi colega el profesor B.C. Berndt sus amplios e insistentes comentarios y conecciones. También le estoy agradecido a Carolyn J. Bloemker por su paciencia y su esmerado trabajo de mecanografía del manuscrito revisado, que se llevó a cabo en todo tipo de circunstancias. Por último, agradezco la ayuda y cooperación del personal de Wiley.

Urbana-Champaign, Illinois **Robert G. Bartle**

Contenido

i012 (1503x2451x2 pcx)

i013 (1444x2493x2 pcx)

INTRODUCCION: UN REPASO A LA TEORIA DE CONJUNTOS

El concepto de conjunto es fundamental para las matemáticas ya que todos los objetos y construcciones matemáticos se originan en la teoría de conjuntos. Debido a la importancia fundamental de ésta, aquí se presenta un breve resumen de las nociones de la teoría de conjuntos que se emplean con frecuencia en este texto. Sin embargo, dado que el propósito de este libro es presentar los *elementos* (más que los *fundamentos*) del análisis real, se adoptará un enfoque bastante pragmático y sencillo. Será suficiente un estudio informal y se dará por entendido el término "conjunto" considerándolo como sinónimo de las palabras "clase", "colección", "agrupación" y "conglomerado". No se pretende definir estos términos ni ofrecer una lista de axiomas para la teoría de conjuntos. El lector que tenga un grado de conocimientos tal que le resulte incómodo este desarrollo informal, deberá consultar la bibliografía de teoría de conjuntos que aparece al final del texto. En ella podrá encontrar la forma en que es posible fundamentar este material sobre bases axiomáticas. Se dará cuenta de que esta axiomatización es un desarrollo muy interesante en los fundamentos de las matemáticas. Sin embargo, puesto que se le considera fuera del área temática de la presente obra, no se entrará aquí en detalles.

Se recomienda leer esta introducción rápidamente a fin de familiarizarse con las notaciones que se emplean. A diferencia de los demás capítulos, los cuales se deben *estudiar*, esta introducción puede considerarse como material de referencia y no se le debe dedicar mucho tiempo.

Sección 1 Algebra de conjuntos

Si A denota un conjunto y si x es un elemento, a menudo es conveniente escribir

$$x \in A$$

como abreviatura de la afirmación de que x es un **elemento de** A, o que x es un **miembro** del conjunto A, o que el conjunto A **contiene al elemento** x, o que x

está en A. No se analizará más la naturaleza de esta propiedad, o sea la de ser un elemento de un conjunto. Para la mayoría de los propósitos es posible emplear el significado intuitivo de "pertenencia" y resulta innecesaria una caracterización axiomática de esta relación.

Si A es un conjunto y x es un elemento que *no* pertenece a A, a menudo se escribe

$$x \notin A.$$

De acuerdo con el concepto intuitivo de conjunto, se requerirá que exactamente una de las dos posibilidades

$$x \in A, \qquad x \notin A,$$

se cumpla para un elemento x y un conjunto A.

Si A y B son dos conjuntos y x es un elemento, entonces en principio existen cuatro posibilidades (Fig. 1.1):

(1) $x \in A$	y	$x \in B$;	(2) $x \in A$	y	$x \notin B$;
(3) $x \notin A$	y	$x \in B$;	(4) $x \notin A$	y	$x \notin B$.

Si el segundo caso no sucede (es decir, si todo elemento de A también es un elemento de B), entonces se dice que A está **contenido** en B, o que B **contiene** a A, o que es un **subconjunto** de B, y se escribe

$$A \subseteq B \qquad \text{ó} \qquad B \supseteq A.$$

Si $A \subseteq B$ y existe algún elemento en B que no esté en A, se dice que A es un **subconjunto propio** de B.

Es importante observar que la afirmación de que $A \subseteq B$ no excluye automáticamente la posibilidad de que A cubra todo B. Cuando esto ocurre los conjuntos A y B son "iguales" en el sentido que a continuación se define.

1.1 DEFINICION. Dos conjuntos son **iguales** si contienen los mismos elementos. Si los conjuntos A y B son iguales, se escribe $A = B$.

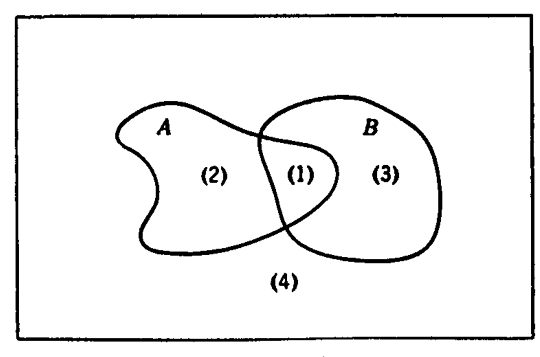


Figura 1.1

Así, para poder probar que los conjuntos A y B son iguales se debe probar que no pueden ocurrir las posibilidades (2) y (3) antes mencionadas. Igualmente, se debe probar que tanto $A \subseteq B$ como $B \subseteq A$.

No es fácil definir con precisión la palabra "propiedad". Sin embargo, se usará de la manera más usual (la informal). Si P designa una propiedad que sea significativa para una colección de elementos, entonces se escribirá

$$\{x : P(x)\}$$

para designar al conjunto de todos los elementos x para los cuales se cumple la propiedad P. Por lo general esto se lee como "el conjunto de todas las x tales que $P(x)$". A menudo es conveniente especificar cuáles elementos se están probando para la propiedad P. Entonces con frecuencia se escribirá

$$\{x \in S : P(x)\}$$

para designar al subconjunto de S para el cual se cumple la propiedad P.

EJEMPLOS. *(a)* Si $\mathbf{N} = \{1, 2, 3, \ldots\}$ designa al conjunto de números naturales, entonces el conjunto

$$\{x \in \mathbf{N} : x^2 - 3x + 2 = 0\}$$

consta de aquellos números naturales que satisfacen la ecuación dada. Las únicas soluciones de la ecuación cuadrática $x^2 - 3x + 2 = 0$ son $x = 1$ y $x = 2$. Por lo tanto, en vez de escribir la expresión anterior (puesto que se tiene información detallada con respecto a todos los elementos del conjunto considerado) por lo común se denota a este conjunto como $\{1, 2\}$ es decir, enumerando los elementos del conjunto.

(b) Algunas veces se puede usar una fórmula para abreviar la descripción de un conjunto. Por ejemplo, el conjunto de todos los números naturales pares podría denotarse por $\{2x : x \in \mathbf{N}\}$, en vez de la expresión $\{y \in \mathbf{N} : y = 2x, x \in \mathbf{N}\}$.

(c) El conjunto $\{x \in \mathbf{N} : 6 < x < 9\}$ se puede escribir explícitamente como $\{7, 8\}$, es decir, mostrando los elementos del conjunto. Desde luego existen muchas otras descripciones posibles para este conjunto. Por ejemplo:

$$\{x \in \mathbf{N} : 40 < x^2 < 80\},$$
$$\{x \in \mathbf{N} : x^2 - 15x + 56 = 0\},$$
$$\{7 + x : x = 0 \quad \text{ó} \quad x = 1\}.$$

(d) Además del conjunto de **números naturales** (que consta de los elementos designados con 1,2,3,...), al que sistemáticamente se designará aquí con **N**, existen algunos otros conjuntos cuya notación convencional se habrá de presentar. El conjunto de los **enteros** es

$$\mathbf{Z} = \{0, 1, -1, 2, -2, 3, -3, \ldots\}.$$

El conjunto de **números racionales** es

$$\boldsymbol{Q} = \{m/n : m, n \in \boldsymbol{Z} \quad \text{y} \quad n \neq 0\}.$$

Se estudiarán los conjuntos $\boldsymbol{N}$, $\boldsymbol{Z}$, y $\boldsymbol{Q}$ suponiendo que existe una comprensión adecuada de los mismos por parte del lector y no se reexaminarán sus propiedades en forma detallada. El conjunto $\boldsymbol{R}$ de los números reales, es de fundamental importancia para los estudios posteriores y se analizará en las secciones 4 - 6. Un subconjunto específico de $\boldsymbol{R}$ que será útil es el **intervalo unitario**

$$\boldsymbol{I} = \{x \in \boldsymbol{R} : 0 \leq x \leq 1\}.$$

Por último, se denotará al conjunto de **números complejos** mediante la letra $\boldsymbol{C}$. En la sección 13 se da una definición más detallada y una breve descripción de algunas de sus propiedades.

Operaciones con conjuntos

A continuación se dan algunos métodos para construir conjuntos a partir de otros ya dados.

1.2 DEFINICION. Si A y B son conjuntos, entonces su **intersección** es el conjunto de todos los elementos que pertenecen tanto a A como a B. La intersección de los conjuntos A, y B se denotan mediante el símbolo $A \cap B$, que se lee "A intersección B". (Fig. 1.2)

1.3 DEFINICION. Si A y B son conjuntos, entonces su **unión** es el conjunto de elementos que pertenecen a A o a B, o tanto a A como a B. La unión de los conjuntos A y B se denotan mediante el símbolo $A \cup B$, que se lee "A unión B". (Véase la Fig. 1.3.)

También se podría definir $A \cap B$ y $A \cup B$ como

$$A \cap B = \{x : x \in A \quad \text{y} \quad x \in B\},$$
$$A \cup B = \{x : x \in A \quad \text{ó} \quad x \in B\}.$$

Con respecto a esto último, es importante observar que la palabra "o" se está empleando en el sentido de inclusión, que es usual en matemáticas y lógica. En la terminología legal este sentido de inclusión algunas veces se indica mediante "y/o".

Se ha supuesto tácitamente que la intersección y la unión de dos conjuntos es nuevamente un conjunto. Entre otras cosas, esto exige que debe existir un conjunto que no contenga ningún elemento (ya que si A y B no tienen elementos en común, su intersección no tiene elementos).

1.4 DEFINICION. Al conjunto que no contiene elementos se le llama conjunto **vacío** o **vacuo** y se denota mediante el símbolo $\emptyset$. Si A y B

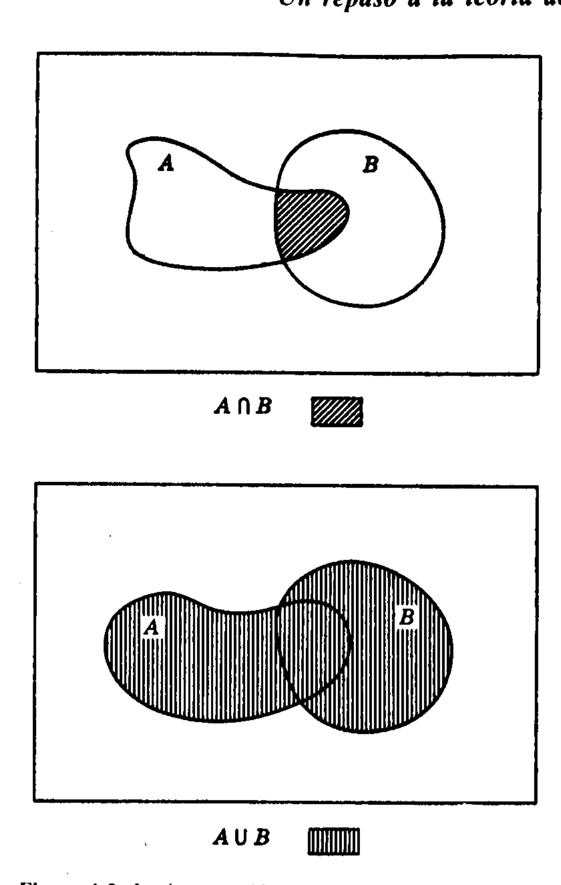


Figura 1.2. La intersección y unión de dos conjuntos.

son conjuntos que no tienen elementos en común (es decirn, si $A \cap B = \emptyset$), entonces se dice que A y B son **ajenos** o que no son **intersecables.**

El siguiente resultado da algunas de las propiedades algebraicas de las operaciones con conjuntos antes definidas. Puesto que las demostraciones de estas afirmaciones son rutinarias, se dejará la mayor parte de ellas como ejercicios para el lector.

1.5 TEOREMA. Sean *A, B, C cualesquier conjuntos;* entonces

(a) $A \cap A = A, \quad A \cup A = A;$
(b) $A \cap B = B \cap A, \quad A \cup B = B \cup A;$
(c) $(A \cap B) \cap C = A \cap (B \cap C), \quad (A \cup B) \cup C = A \cup (B \cup C);$
(d) $A \cap (B \cup C) = (A \cap B) \cup (A \cap C),$
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$

En ocasiones se hace referencia a estas igualdades como las propiedades: de *igualdad, conmutativa, asociativa* y *distributiva,* respectivamente, de las operaciones de intersección y unión de conjuntos.

Como muestra de las demostraciones, se probará la primera ecuación del inciso *(d)*. Sea x un elemento de $A \cap (B \cup C)$, entonces $x \in A$ y $x \in B \cup C$. Esto significa que $x \in A$ y que $x \in B$ o $x \in C$. De modo que se puede tener (i) $x \in A$ y $x \in B$, o bien (ii) $x \in A$ y $x \in C$. Por lo tanto, $x \in A \cap B$ o $x \in A \cap C$, por lo que $x \in (A \cap B) \cup (A \cap C)$.

A la inversa sea y un elemento de $(A \cap B) \cup (A \cap C)$. Entonces (iii) $y \in A \cap B$, o (iv) $y \in A \cap C$. de donde se infiere que $y \in A$, y que $y \in B$ o $y \in C$. Por lo tanto, $y \in A$ y $y \in B \cup C$ de modo que $y \in A \cap (B \cup C)$. En consecuencia, $(A \cap B) \cup (A \cap C)$ es un subconjunto de $A \cap (B \cup C)$. Con base en la definición 1.1 se concluye que los conjuntos $A \cap (B \cup C)$ y $(A \cap B) \cup (A \cap C)$ son iguales.

Como una indicación de un método alternativo, se hace notar que en principio existen, un total de $8(=2^3)$ posibilidades para un elemento x en relación con tres conjuntos A, B, C (Fig. 1.3) a saber:

(1) $x \in A$, $x \in B$, $x \in C$;	(2) $x \in A$, $x \in B$, $x \notin C$;
(3) $x \in A$, $x \notin B$, $x \in C$;	(4) $x \in A$, $x \notin B$, $x \notin C$;
(5) $x \notin A$, $x \in B$, $x \in C$;	(6) $x \notin A$, $x \in B$, $x \notin C$;
(7) $x \notin A$, $x \notin B$, $x \in C$;	(8) $x \notin A$, $x \notin B$, $x \notin C$.

La demostración consiste en probar que ambos miembros de la primera ecuación en *(d)* contienen aquellos y sólo aquellos elementos x comprendidos en los casos (1), (2) o (3).

En vista de las relaciones establecidas en el teorema 1.5 *(c)*, por lo general se eliminan los paréntesis y simplemente se escribe

$$A \cap B \cap C, \qquad A \cup B \cup C.$$

Es posible demostrar que si $\{A_1, A_2, \ldots, A_n\}$ es una colección de conjuntos, entonces existe un conjunto A definido de manera única que consta de todos los elementos que pertenecen *cuando menos a uno* de los conjuntos A_j, $j = 1, 2, \ldots, n$; y existe un conjunto B definido de manera única que

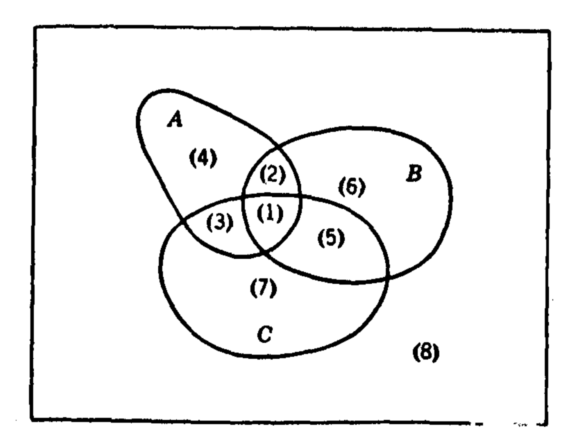


Figura 1.3.

i023 (1528x2468x2 pcx)

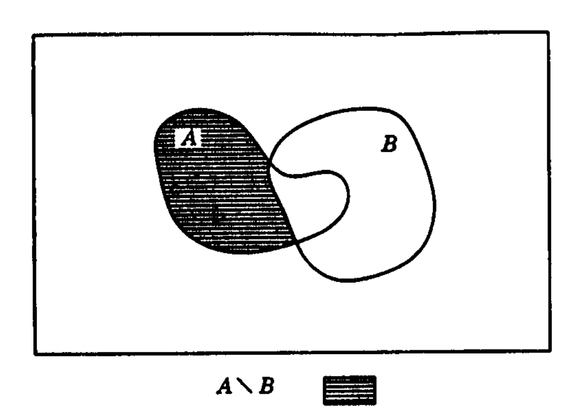


Figura 1.4. El complemento relativo.

consta de todos los elementos que pertenecen a todos los conjuntos A_j, $j = 1, 2, \ldots, n$. Eliminando el uso de los paréntesis, se escribe

$$A = A_1 \cup A_2 \cup \cdots \cup A_n, \qquad B = A_1 \cap A_2 \cap \cdots \cap A_n.$$

Algunas veces, con el objeto de abreviar, se imita la notación que se utiliza en cálculo para las sumas y se emplea una notación más condensada, como

$$A = \bigcup_{j=1}^{n} A_j = \bigcup\{A_j : j = 1, 2, \ldots, n\},$$

$$B = \bigcap_{j=1}^{n} A_j = \bigcap\{A_j : j = 1, 2, \ldots, n\}.$$

En forma análoga, si para cada j en un conjunto J hay un conjunto A_j, entonces $\bigcup\{A_j : j \in J\}$ denota al conjunto de todos los elementos que pertenecen *cuando menos a uno* de los conjuntos A_j. De la misma manera, $\bigcap\{A_j : j \in J\}$ denota al conjunto de todos los elementos que pertenecen a *todos* los conjuntos A_j para $j \in J$.

A continuación se presenta otro método para construir un conjunto nuevo a partir de dos conjuntos dados.

1.6. DEFINICION. Si A y B son conjuntos, entonces el complemento de B respecto de A es el conjunto de todos los elementos de A que no pertenecen a B. Se denotará este conjunto con $A \setminus B$ (se lee "A menos B"), aunque las notaciones afines $A - B$ y $A \sim B$ son utilizadas en ocasiones por otros autores. (Fig. 1.4.)

Con la notación que se acaba de introducir se tiene

$$A \setminus B = \{x \in A : x \notin B\}.$$

A veces el conjunto A se sobreentiende y no es necesario mencionarlo explícitamente. En este caso simplemente se hará referencia al *complemento* de B y se denotará a $A \setminus B$ con $\mathscr{C}(B)$.

Volviendo a la Fig. 1.1, se puede ver que los elementos x que satisfacen (1) pertenecen a $A \cap B$; aquellos que satisfacen (2) pertenecen a $A \setminus B$; y los que satisfacen (3) pertenecen a $B \setminus A$. Ahora se demostrará que A es la unión de los conjuntos $A \cap B$ y $A \setminus B$.

1.7 TEOREMA. *Los conjuntos $A \cap B$ y $A \setminus B$ son no intersecables y*

$$A = (A \cap B) \cup (A \setminus B).$$

DEMOSTRACION. Suponga que $x \in A \cap B$ y $x \in A \setminus B$. Lo segundo afirma que $x \in A$ y $x \notin B$ lo cual contradice la relación $x \in A \cap B$. Por lo tanto, los conjuntos son ajenos.

Si $x \in A$, entonces $x \in B$ o $x \notin B$. En el primer caso $x \in A$ y $x \in B$ de manera que $x \in A \cap B$.. En el segundo, $x \in A$ y $x \notin B$ de modo que $x \in A \setminus B$. Esto prueba que A es un subconjunto de $(A \cap B) \cup (A \setminus B)$. De manera inversa, si $y \in (A \cap B) \cup (A \setminus B$, entonces $y \in A \cap B$, o bien $y \in A \setminus B$. En cualquiera de estos casos se tiene $y \in A$ demostrándose que $(A \cap B) \cup (A \setminus B)$ es un subconjunto de A.

Se enunciarán ahora las leyes de *De Morgan*† para tres conjuntos; una formulación más general se presentará en los ejercicios.

1.8 TEOREMA. *Si A, B, C, son cualesquier conjuntos, entonces*

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C),$$
$$A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

DEMOSTRACION. Se hará la demostración de la primera relación dejando la segunda al lector. Para probar la igualdad de los conjuntos se demuestra que todo elemento en $A \setminus (B \cup C)$ está contenido tanto en $(A \setminus B)$ como en $(A \setminus C)$ e inversamente.

Si x está en $A \setminus (B \cup C)$, entonces x está en A pero no está en $B \cup C$. De modo que x está en A, pero no está ni en B ni en C. (¿por qué?) Por lo tanto, x está en A pero no en B, y x está en A pero no en C. Es decir, $x \in A \setminus B$ y $x \in A \setminus C$, demostrándose que $x \in (A \setminus B) \cap (A \setminus C)$.

Al contrario, si $x \in (A \setminus B) \cap (A \setminus C)$, entonces $x \in (A \setminus B)$ y $x \in (A \setminus C)$. De modo que $x \in A$, y además $x \notin B$ y $x \notin C$. Se infiere que $x \in A$ y $x \notin (B \cup C)$, de tal manera que $x \in A \setminus (B \cup C)$.

Puesto que los conjuntos $(A \setminus B) \cap (A \setminus C)$ y $A \setminus (B \cup C)$ contienen los mismos elementos, por la definición 1.1, son iguales.

†AUGUSTUS DE MORGAN (1806-1873) ejerció el magisterio en "University College", Londres. Fue matemático y dialéctico y contribuyó a preparar el camino de la lógica matemática moderna.

Producto cartesiano

Se define ahora el producto cartesiano‡ de dos conjuntos.

1.9. DEFINICION. Si A y B son dos conjuntos no vacíos, entonces el **producto cartesiano** $A \times B$ de A y B es el conjunto de todos los pares ordenados (a, b) con $a \in A$ y $b \in B$. (Fig. 1.5)

(La definición que se acaba de dar es un tanto informal ya que no se ha definido lo que es un "par ordenado". No nos detendremos más en esta cuestión sino para mencionar que el par ordenado (a, b) se podría definir como el conjunto cuyos únicos elementos son $\{a\}, \{a, b\}$. Se puede probar entonces que los pares ordenados (a, b) y (a', b') son iguales si y sólo si $a = a'$ y $b = b'$. Esta es la propiedad fundamental de los pares ordenados.)

De tal modo, si $A = \{1, 2, 3\}$ y $B = \{4, 5\}$, entonces el conjunto $A \times B$ es el conjunto cuyos elementos son los pares ordenados

$$(1, 4), (1, 5), (2, 4), (2, 5), (3, 4), (3, 5).$$

Se puede visualizar al conjunto $A \times B$ como la colección de seis puntos en el plano, los cuales tienen las coordenadas que se acaban de enumerar.

Con frecuencia se dibuja un diagrama (como el de la Fig. 1.5) para ilustrar el producto cartesiano de dos conjuntos A, B. Sin embargo, se debe tomar en cuenta que este diagrama es en cierto modo una simplificación. Por ejemplo, si $A = \{x \in \boldsymbol{R} : 1 \leq x \leq 2\}$ y $B = \{x \in \boldsymbol{R} : 0 \leq x \leq 1$ $2 \leq x \leq 3\}$, entonces en vez de un rectángulo se tendrá un dibujo como el de la Fig. 1.6.

Ejercicios

1.A. Dibuje un diagrama para representar cada uno de los conjuntos mencionados en el teorema 1.5.

1.B. Demuestre la parte *(c)* del teorema 1.5.

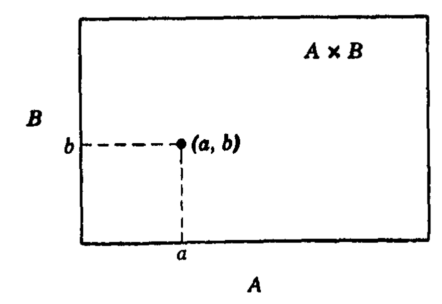


Figura 1.5. El producto cartesiano.

‡RENE DESCARTES (1596-1650), creador de la geometría analítica, fue un gentil hombre francés, soldado, matemático y uno de los más grandes filósofos de todos los tiempos.

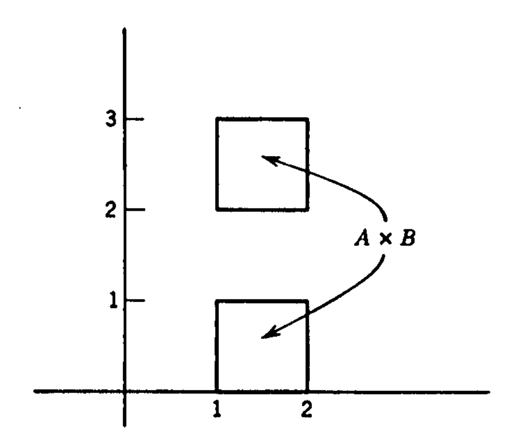


Figura 1.6. El producto cartesiano.

1.C. Demuestre la segunda parte de *(d)* del teorema 1.5.

1.D. Demuestre que $A \subseteq B$ si y sólo si $A \cap B = A$.

1.E. Demuestre que el conjunto D de todos los elementos que pertenecen a A o bien a B pero no a los dos está dado por

$$D = (A \setminus B) \cup (B \setminus A).$$

A este conjunto D a menudo se le llama **diferencia simétrica** de A y B. Represéntelo por medio de un diagrama.

1.F. Demuestre que la diferencia simétrica D que se define en el ejercicio anterior está dada también por $D = (A \cup B) \setminus (A \cap B)$.

1.G. Si $B \subseteq A$, demuestre que $B = A \setminus (A \setminus B)$.

1.H. Si A y B son cualesquier conjuntos, demuestre que $A \cap B = A \setminus (A \setminus B)$.

1.I. Si $\{A_1, A_2, \ldots, A_n\}$ es una colección de conjuntos y E cualquier conjunto, demuestre que

$$E \cap \bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{n} (E \cap A_j), \qquad E \cup \bigcup_{j=1}^{n} A_j = \bigcup_{j=1}^{n} (E \cup A_j).$$

1.J. Si $\{A_1, A_2, \ldots, A_n\}$ es una colección de conjuntos y E cualquier conjunto, demuestre que

$$E \cap \bigcap_{j=1}^{n} A_j = \bigcap_{j=1}^{n} (E \cap A_j), \qquad E \cup \bigcap_{j=1}^{n} A_j = \bigcap_{j=1}^{n} (E \cup A_j).$$

1.K. Sea E un conjunto y $\{A_1, A_2, \ldots, A_n\}$ una colección de conjuntos. Establezca las leyes de De Morgan:

$$E \setminus \bigcap_{j=1}^{n} A_j = \bigcup_{j=1}^{n} (E \setminus A_j), \qquad E \setminus \bigcup_{j=1}^{n} A_j = \bigcap_{j=1}^{n} (E \setminus A_j).$$

Observe que si $E \setminus A_j$ se denota por $\mathscr{C}(A_j)$, estas relaciones toman la forma

$$\mathscr{C}\left(\bigcap_{j=1}^{n} A_j\right) = \bigcup_{j=1}^{n} \mathscr{C}(A_j), \qquad \mathscr{C}\left(\bigcup_{j=1}^{n} A_j\right) = \bigcap_{j=1}^{n} \mathscr{C}(A_j).$$

1.L. Sea J cualquier conjunto y suponga A_j contenido en X para cada $j \in J$. Demuestre que

$$\mathscr{C}(\bigcap\{A_j : j \in J\}) = \bigcup\{\mathscr{C}(A_j) : j \in J\},$$
$$\mathscr{C}(\bigcup\{A_j : j \in J\}) = \bigcap\{\mathscr{C}(A_j) : j \in J\}.$$

1.M. Si B_1 y B_2 son subconjuntos de B y si $B = B_1 \cup B_2$, entonces

$$A \times B = (A \times B_1) \cup (A \times B_2).$$

Sección 2 Funciones

Ahora se estudiará la noción fundamental de *función* o *aplicación*. Se observará que una función es un tipo especial de conjunto, aunque existen otras acepciones que con frecuencia son más aceptables. En las secciones siguientes se estudiarán varios tipos de funciones, pero por lo general serán de naturaleza menos abstracta que la que se estudia en esta sección introductoria.

Para el matemático de hace un siglo el término "función" por lo común significaba una fórmula determinada como

$$f(x) = x^2 + 3x - 5,$$

la cual asocia a cada número real x otro número real $f(x)$. El hecho de que ciertas fórmulas, tales como

$$g(x) = \sqrt{x - 5},$$

no den origen a números reales para todo valor real de x era, desde luego, bien sabido, pero no se consideraba que fuera razón suficiente como para hacer necesaria una extensión del concepto de función. Tal vez se podría suscitar una controversia entre aquellos matemáticos en cuanto a si el valor absoluto

$$h(x) = |x|$$

de un número real es una "función honesta" o no, ya que, después de todo, la definición de $|x|$ está dada "en partes" por

$$|x| = \begin{cases} x, & \text{si} \quad x \geq 0, \\ -x, & \text{si} \quad x < 0. \end{cases}$$

Conforme se desarrollaron las matemáticas, resultó cada vez más claro que el requisito de que una función fuera una fórmula era indebidamente restrictivo y que una definición más general sería útil. También resultó evidente que es importante hacer una clara distinción entre la función misma y los valores de la función. Quizá el lector por causas que le son ajenas se en-

cuentre, en la misma situación que los matemáticos de hace un siglo con respecto a estas dos consideraciones. Aquí se pretende ponerle al tanto de la conceptualización actual, pero ello se hará en dos pasos. La primera definición revisada de una función sería:

Una función f de un conjunto A a un conjunto B es una regla de correspondencia que asigna a cada x en un cierto subconjunto D de A, un elemento determinado de manera única $f(x)$ de B.

Desde luego, las fórmulas explícitas del tipo que se mencionó antes están incluidas en esta definición tentativa. La definición propuesta admite la posibilidad de que la función pueda no estar definida para ciertos elementos de A y además admite la consideración de funciones para las cuales los conjuntos A y B no son necesariamente números reales (sino que incluso pueden ser escritorios y sillas, o gatos y perros).

A pesar de lo aceptable que pueda parecer la definición propuesta, tiene un gran defecto: no es clara. Aún queda el problema de interpretar la frase "regla de correspondencia". Sin duda el lector puede pensar en frases que le satisfagan más que la anterior; sin embargo, no es fácil que pueda disipar del todo la confusión. La solución más satisfactoria parece ser la de definir "función" íntegramente en términos de conjuntos y de las propiedades que se introdujeron en la sección anterior. Esto tiene la desventaja de ser más artificial y de perder algo del contenido intuitivo de la descripción antes dada; sin embargo, la claridad que se obtiene pesa más que estas desventajas.

La idea clave es pensar en la gráfica de la función, o sea en una colección de pares ordenados. Se hace la aclaración de que una colección arbitraria de pares ordenados no puede ser la gráfica de una función, ya que una vez que se hace referencia al primer miembro del par ordenado el segundo queda determinado de manera única.

2.1. DEFINICION. Sean A y B conjuntos (no necesariamente distintos). Una **función de** A **a** B es un conjunto f de pares ordenados en $A \times B$ con la propiedad de que si (a, b) son elementos de f, entonces $(a{=}b')$ Al conjunto de todos los elementos de A que pueden aparecer como primeros miembros de elementos de f se le llama **dominio** de f y se denotará por $D(f)$. Al conjunto de todos los elementos de B que puedan aparecer como segundos miembros de elementos de f se le llama **rango** de f (o **conjunto de valores** de f) y se denotará por $R(f)$. En el caso en que $D(f) = A$, con frecuencia se dice que f **mapea** A **en** B (o es un **mapeo,** de A en B) y se escribe $f: A \to B$.

Si (a, b) es un elemento de una función f, entonces es común escribir

$$b = f(a) \qquad \text{ó} \qquad f: a \mapsto b$$

en vez de $(a, b) \in f$. Por lo general se hace referencia al elemento b como el **valor** de f en el punto a, o la **imagen bajo** f del punto a.

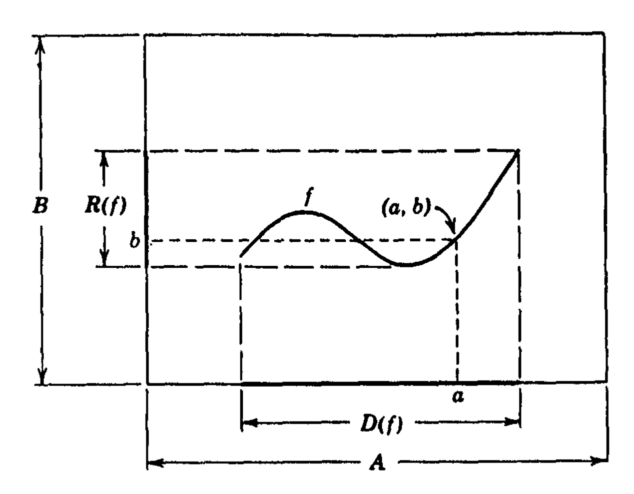


Figura 2.1. Una función vista como gráfica.

Representación tabular.

Una manera de visualizar una función es por medio de una gráfica. Otro método importante y que se utiliza mucho es por medio de una *tabla.* Considere la tabla 2.1, que se podría encontrar en la sección deportiva del *Foosland Bugle-Gazette.*

El dominio de esta función f de tiro libre consta de los nueve jugadores

$D(f)$ = {Anderson, Bade, Bateman, Hochschild, Kakutani, Kovalevsky, Osborn, Peressini, Rosenberg},

mientras que el rango de la función consta de los seis números

$$R(f) = \{0, 1, 2, 4, 5, 8\}.$$

Los elementos de la función propiamente dichos son los pares ordenados

(Anderson, 2), (Bade, 0), (Bateman, 5),
(Hochschild, 1), (Kakutani, 4), (Kovalevsky, 8),
(Osborn, 0), (Peressini, 2), (Rosenberg, 4).

En representaciones tabulares de este tipo por lo regular sólo se escribe el dominio de la función en la columna del lado izquierdo, (puesto que no hay necesidad de mencionar a los miembros del equipo que no jugaron). Se puede decir que el valor de esta función f de tiro libre en Anderson es 2 y escribir $f(\text{Anderson}) = 2$, o Anderson $\mapsto 2$, y así sucesivamente.

i030 (1528x2476x2 pcx)

Tabla 2.1

Jugador	Tiros libres efectuados
Anderson	2
Bade	0
Bateman	5
Hochschild	1
Kakutani	4
Kovalevsky	8
Osborn	0
Peressini	2
Rosenberg	4

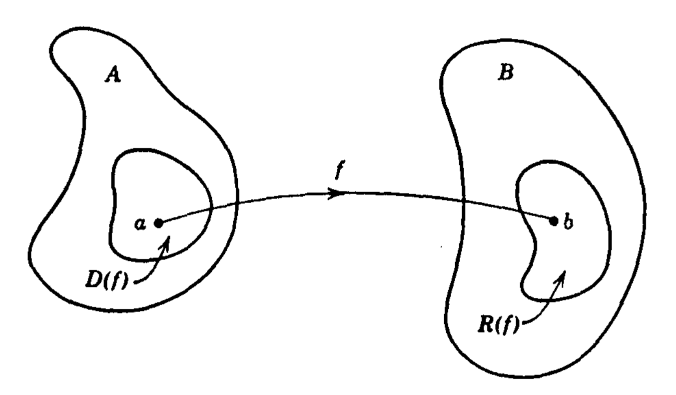


Figura 2.2. Una función vista como transformación.

Todos estamos familiarizados con el uso de tablas para transmitir información. Estas son ejemplos importantes de funciones y por lo general son de tal índole que sería difícil representarlas en términos de una fórmula.

Transformaciones y máquinas

Existe otra manera de visualizar una función: como una *transformación* de una parte del conjunto A hacia una parte de B. En estos términos, cuando $(a, b) \in f$, nos imaginamos a f como si tomara el elemento a del subconjunto $D(f)$ de A y lo "transformara" o "mapeara" en un elemento $b = f(a)$ del subconjunto $R(f)$ de B. A menudo se dibuja un diagrama como el de la Fig. 2.2. Con frecuencia se usa esta representación geométrica de una función aun cuando los conjuntos A y B no sean subconjuntos del plano.

Existe otra manera de visualizar una función: como una *máquina* que acepta elementos de $D(f)$ como materia prima produciendo elementos correspondientes de $R(f)$ como producto final. Si se toma un elemento x de $D(f)$ y se

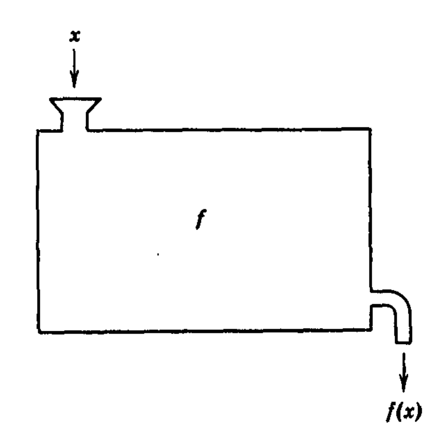


Figura 2.3. Una función vista como máquina.

introduce en f, saldrá el valor correspondiente $f(x)$. Si se mete un elemento distinto y de $D(f)$ en f, se obtiene $f(y)$ (que puede o no diferir de $f(x)$). Si se trata de introducir en f algo que no pertenezca a $D(f)$, se verá que no es aceptado, ya que f sólo puede operar con elementos pertenecientes a $D(f)$. (Fig. 2.3.)

Esta última representación aclara la diferencia entre f y $f(x)$; lo primero es la máquina y lo segundo es el rendimiento de la máquina al ser alimentada con x. Desde luego, es importante distinguir entre una máquina y sus productos. Sólo un tonto confundiría un molino de carne con la carne molida; sin embargo, tantas personas han confundido algunas funciones con sus valores que vale la pena hacer un esfuerzo para distinguir entre los dos conceptos por medio de su notación.

Restricciones y extensiones de funciones

Si f es una función con dominio $D(f)$ y D_1 es un subconjunto de $D(f)$, con frecuencia resulta útil definir una nueva función f_1 con dominio D_1 por medio de $f_1(x) = f(x)$ para toda $x \in D_1$. Esta función f_1 se llama la *restricción de f al conjunto D_1*. En términos de la definición 2.2, se tiene

$$f_1 = \{(a, b) \in f : a \in D_1\}.$$

Algunas veces se escribe $f_1 = f \mid D_1$ para denotar la restricción de la función f al conjunto D_1.

Una estructura análoga (que resulta menos artificial) es el concepto de "extensión". Si g es una función con dominio $D(g)$ y $D_2 \supseteq D(g)$, entonces cualquier función g_2 con dominio D_2 tal que $g_2(x) = g(x)$ para toda $x \in D(g)$ se llama una *extensión de g al conjunto D_2*.

Composición de funciones.

Ahora se intenta "componer" dos funciones aplicando primero f a cada x en $D(f)$ y después aplicando g a $f(x)$ siempre que sea posible (es decir, cuando $f(x)$ pertenezca a $D(g)$. Al hacer esto se debe tener cierto cuidado en lo que respecta al dominio de la función resultante. Por ejemplo, si f está definida en $\mathbf{R}$ por $f(x) = x^3$ y si g está definida para $x \geq 0$ por $g(x) = \sqrt{x}$, entonces la composición $g \circ f$ sólo se puede definir para $x \geq 0$, y para estos números reales deberá tener el valor $\sqrt{x^3}$.

2.2 DEFINICION. Sea f una función con dominio $D(f)$ en A y rango $R(f)$ en B y sea g una función con dominio $D(g)$ en B y rango $R(g)$ en C. (Fig. 2.4) La **composición** $g \circ f$ (observe el orden) es la función desde A a C dada por

$$g \circ f = \{(a, c) \in A \times C\text{: existe un elemento } b \in B \text{ tal que } (a, b) \in f \text{ y } (b, c) \in g\}.$$

2.3 TEOREMA. *Si f y g son funciones, la composición $g \circ f$ es una función con*

$$D(g \circ f) = \{x \in D(f) : f(x) \in D(g)\},$$
$$R(g \circ f) = \{g(f(x)) : x \in D(g \circ f)\}.$$

2.4 EJEMPLOS. *(a)* Sean f y g funciones cuyos valores en el número real x son los números reales dados por†

$$f(x) = 2x, \qquad g(x) = 3x^2 - 1.$$

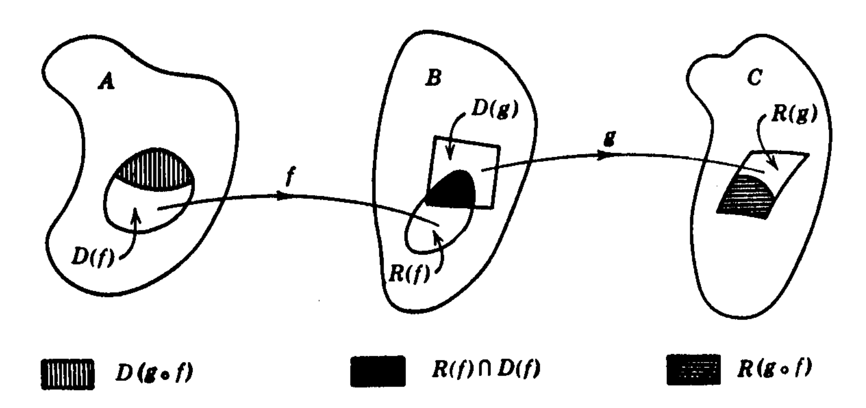


Figura 2.4. Composición de funciones.

† lo también se denota escribiendo $f : x \mapsto 2x$ y $g : x \mapsto 3x^2 - 1$ para $x \in \mathbf{R}$.

Puesto que $D(g)$ es el conjunto $\mathbf{R}$ de todos los números reales y $\mathbf{R}(f) \subseteq D(g)$, el dominio $D(g \circ f)$ también es $\mathbf{R}$ y $g \circ f(x) = 3(2x)^2 - 1 = 12x^2 - 1$. Por otro lado, $D(f \circ g) = \mathbf{R}$, pero $f \circ g(x) = 2(3x^2 - 1) = 6x^2 - 2$.

(b) Si h es la función con $D(h) = \{x \in \mathbf{R} : x \geq 1\}$ definida por

$$h(x) = \sqrt{x - 1},$$

y si f es como en el inciso *(a)*, entonces $D(h \circ f) = \{x \in \mathbf{R} : 2x \geq 1\} = \{x \in \mathbf{R} : x \geq \frac{1}{2}\}$ y $h \circ f(x) = \sqrt{2x - 1}$. Además $D(f \circ h) = \{x \in \mathbf{R} : x \geq 1\}$ y $f \circ h(x) = 2\sqrt{x - 1}$. Si g es la función del inciso *(a)*, entonces $D(h \circ g) = \{x \in \mathbf{R} : 3x^2 - 1 \geq 1\} = x \in \mathbf{R} : x \leq -\sqrt{\frac{2}{3}}$ ó $x \geq \sqrt{\frac{2}{3}}\}$ y $h \circ g(x) = \sqrt{3x^2 - 2}$. Además $D(g \circ h) = \{x \in \mathbf{R} : x \geq 1\}$ y $g \circ h(x) = 3x - 4$. (Observe que la fórmula que representa a $g \circ h$ tiene significado para otros valores de x distintos de aquellos en el dominio de $g \circ h$.)

(c) Sean F, G funciones con dominios $D(F) = \{x \in \mathbf{R} : x \geq 0\}$, y $D(G) = \mathbf{R}$, tales que los valores de F y G en un punto x de sus dominios sean

$$F(x) = \sqrt{x}, \qquad G(x) = -x^2 - 1.$$

Entonces $D(G \circ F) = \{x \in \mathbf{R} : x \geq 0\}$ y $G \circ F(x) = -x - 1$, mientras que $D(F \circ G) = \{x \in D(G) : G(x) \in D(f)\}$. Este último conjunto es vacío ya que $G(x) < 0$ para toda $x \in D(G)$. Por lo tanto, la función $F \circ G$ no está definida en ningún punto, entonces $F \circ G$ es la "función nula".

Funciones inyectivas e inversas

Se dará ahora un método para construir una nueva función a partir de otra dada, siempre que la función original no tome el mismo valor dos veces.

2.5 DEFINICION. Sea f una función con dominio $D(f)$ en A y rango $R(f)$ en B. Se dice que f es **inyectiva** o **uno a uno** si, cada vez que (a, b) y (a', b) son elementos de f, entonces $a = a'$. Si f es inyectiva se puede decir que f es una **inyección.**

En otras palabras, f es inyectiva si y sólo si las dos relaciones $f(a) = b$, $f(a') = b$ implican $a = a'$. Alternativamente f es inyectiva si y sólo si a, a' están en $D(f)$ y $a \neq a'$, entonces $f(a) \neq f(a')$.

Se sostiene que si f es inyectiva desde A a B, entonces el conjunto de pares ordenados en $B \times A$ que se obtienen al intercambiar las componentes de cada uno de los pares ordenados de f da una función g que también es inyectiva.

Se omite la demostración de esta afirmación dejándola como ejercicio; es una buena prueba para el lector. Las relaciones entre f y g son:

$$D(g) = R(f), \qquad R(g) = D(f),$$
$$(a, b) \in f \quad \text{si y sólo si} \quad (b, a) \in g.$$

Esta última afirmación se puede escribir en la forma más usual:

$$b = f(a) \quad \text{si y sólo si} \quad a = g(b).$$

2.6 DEFINICION. Sea f una inyección con dominio $D(f)$ en A y rango $R(f)$ en B. Si $g = \{(b, a) \in B \times A : (a, b) \in f\}$, entonces g es una inyección con dominio $D(g) = R(f)$ en B y con rango $R(g) = D(f)$ en A. La función g se llama función **inversa** de f y se denota por f^{-1}.

La función inversa se puede interpretar desde el punto de vista de un mapeo. (Fig. 2.5) Si f es inyectiva, mapea distintos elementos de $D(f)$ hacia distintos elementos de $R(f)$. De tal manera, cada elemento b de $R(f)$ es la imagen bajo f de un único elemento a en $D(f)$. La función inversa f^{-1} mapea el elemento b hacia este elemento único a.

2.7 EJEMPLOS. *(a)* Sea $F: x \mapsto x^2$ la función con dominio $D(F) = \mathbf{R}$, el conjunto de los número reales, y rango en $\mathbf{R}$ tal que el valor de F en el número real x sea $F(x) = x^2$. (En otras palabras, F es la función $\{(x. x^2) : x \in \mathbf{R}\}$.) Fácilmente se ve que F no es uno a uno; de hecho, ambos pares ordenados, $(2, 4)$, $(-2, 4)$, pertenecen a F. Puesto que F no es uno a uno, no tiene inverso.

(b) Sea f la función con dominio $D(f) = \{x \in \mathbf{R} : x \geq 0\}$ y $R(f) = \mathbf{R}$ cuyo valor para x en $D(f)$ es $D(f)$ es $f(x) = x^2$. Observe que f es la restricción a $D(f)$ de la función F del inciso *(a)*. En términos de pares ordenados, $f = \{(x, x^2) : x \in \mathbf{R}, x \geq 0\}$. A diferencia de la función F del inciso *(a)*, f es inyectiva, ya que si $x^2 = y^2$ con x, y en $D(f)$ entonces $x = y$. (¿por qué?) Por lo tanto, f tiene una función inversa g con $D(g) = R(f) = \{x \in \mathbf{R} : x \geq 0\}$ y $R(g) = D(f) = \{x \in \mathbf{R} : x \geq 0\}$ Además, $y = x^2 = f(x)$ si y sólo si $x = g(y)$. A esta función inversa g por lo general se le llama **función raíz cuadrada positiva** y se denota por

$$g(y) = \sqrt{y}, \qquad y \in \mathbf{R}, \quad y \geq 0.$$

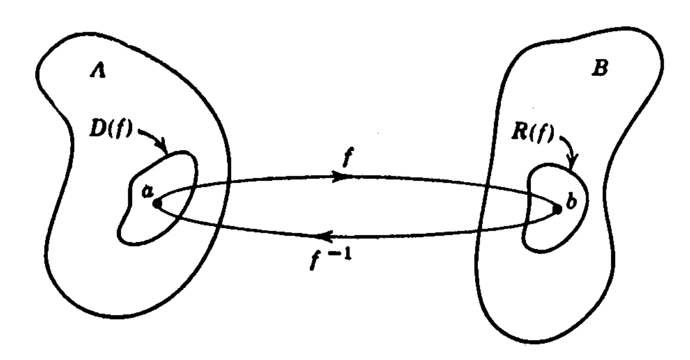


Figura 2.5. La función inversa.

(c) Si f_1 es la función $\{(x, x^2) : x \in \boldsymbol{R}, x \leq 0\}$, entonces, igual que en *(b)*, f_1 es uno a uno y tiene dominio $D(f_1) = \{x \in \boldsymbol{R} : x \leq 0\}$ y como rango $R(f_1) = \{x \in \boldsymbol{R} : x \geq 0\}$. Note que f_1 es la restricción a $D(f_1)$ de la función F del inciso *(a)*. La función g_1 inversa a f se llama la **función raíz cuadrada negativa** y se denota por

$$g_1(y) = -\sqrt{y}, \qquad y \in \boldsymbol{R}, \quad y \geq 0,$$

de tal manera que $g_1(y) \leq 0$.

(d) La función seno F, introducida en trigonometría con $D(F) = \boldsymbol{R}$ y $R(F) = \{y \in \boldsymbol{R} : -1 \leq y \leq +1\}$, se sabe muy bien que no es inyectiva (por ejemplo, sen $0 =$ sen $2\pi = 0$). Sin embargo si se supone que f es su restricción al conjunto $D(f) = \{x \in \boldsymbol{R} : -\pi/2 \leq x \leq +\pi/2\}$, entonces f es inyectiva. Por lo tanto, tiene una función inversa g con $D(g) = R(f)$ y $R(g) = D(f)$. Además, $y =$ sen x con $x \in D(f)$ si y sólo si $x = g(y)$. La función g se llama la (rama principal) de la **función inversa seno** y a menudo se denota por

$$g(y) = \text{Arc sen } y \qquad \text{ó} \qquad g(y) = \text{Sen}^{-1}\, y.$$

Funciones suprayectivas y biyectivas

2.8 DEFINICION. Sea f una función $D(f) \subseteq A$ y rango $R(f) \subseteq B$. Se dice que f es **suprayectiva**, o que f **mapea sobre** B, cuando el rango $R(f) = B$. Si f es suprayectiva, se puede decir que f es una **suprayección.**

Al definir una función, es importante especificar el dominio de la función y el conjunto en el cual se adoptan los valores. Una vez hecho esto se puede averiguar si la función es o no suprayectiva.

2.9 DEFINICION. Una función f con dominio $D(f) \subseteq A$ y rango $R(f) \subseteq B$ se dice que es **biyectiva** si (i) es inyectiva (es decir que es uno a uno), y (ii) es suprayectiva (es decir, aplica $D(f)$ sobre B). Si f es biyectiva, se puede decir que f es una **biyección.**

Imágenes directas e inversas

Sea f una función arbitraria con dominio $D(f)$ en A y rango $R(f)$ en B. *No* se está suponiendo que f sea inyectiva.

2.10 DEFINICION. Si E es un subconjunto de A, entonces la **imagen directa** de E bajo f es el subconjunto de $R(f)$ dado por

$$\{f(x) : x \in E \cap D(f)\}.$$

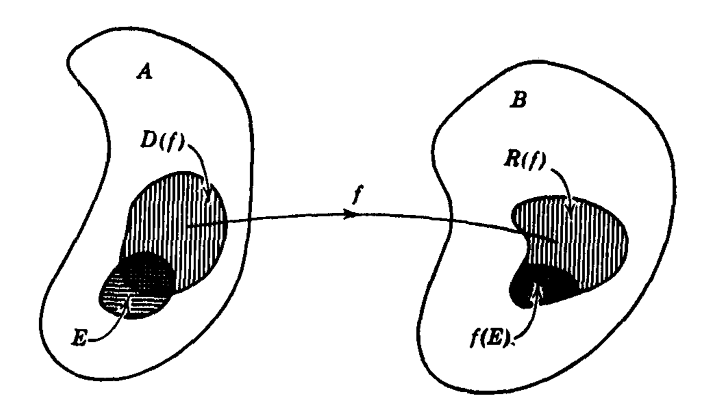


Figura 2.6. Imágenes directas.

Por lo regular se representa la imagen directa de un conjunto E bajo f por medio de la notación $f(E)$. (Fig. 2.6.)

Se observará que si $E \cap D(f) = \emptyset$, entonces $f(E) = \emptyset$. Si E contiene un único punto p de $D(f)$, entonces el conjunto $f(E)$ contiene un único punto $f(p)$. Algunas propiedades de conjuntos se conservan bajo la imagen directa, como se demuestra en seguida.

2.11 TEOREMA. *Sea f una función con dominio en A y rango en B, y sean E, F subconjuntos de A.*

(a) *si* $E \subseteq F$, *entonces* $f(E) \subseteq f(F)$.
(b) $f(E \cap F) \subseteq f(E) \cap f(F)$.
(c) $f(E \cup F) = f(E) \cup f(F)$.
(d) $f(E \setminus F) \subseteq f(E)$.

DEMOSTRACION. (*a*) Si $E \cap D(f)$, entonces $F \cap D(f)$ y de aquí $f(x) \in f(F)$. Puesto que esto es cierto para toda $E \cap D(f)$, se deduce que $f(E) \subseteq f(F)$.

(*b*) Puesto que $E \cap F \subseteq E$, del inciso (*a*) se infiere que $f(E \cap F) \subseteq f(E)$; análogamente, $f(E \cap F) \subseteq f(F)$. Por lo tanto, se concluye que $f(E \cap F) \subseteq f(E) \cap f(F)$.

(*c*) Puesto que $E \subseteq E \cup F$ y $F \subseteq E \cup F$, del inciso (*a*) se infiere que $f(F) \subseteq f(E \cup F)$. Inversamente, si $y \in f(E \cup F)$, entonces existe un elemento $x \in E \cup F$ tal que $y = f(x)$. Dado que $x \in E$ ó $x \in F$, se deduce que $y = f(x) \in f(E)$ o bien que $y \in f(F)$. Por lo que se puede concluir que $f(E \cup F) \subseteq f(E) \cup f(F)$, y esto completa la demostración del inciso (*c*).

(*d*) El inciso (*d*) se infiere directamente de (*a*).

En el ejercicio 2.J se verá que, en general, no es posible reemplazar el signo de inclusión en (*b*) por el de igualdad.

Se introducirá ahora la idea elemental de la imagen inversa de un conjunto bajo una función. Observe que no se pide que la función sea inyectiva.

i037 (1578x2419x2 pcx)

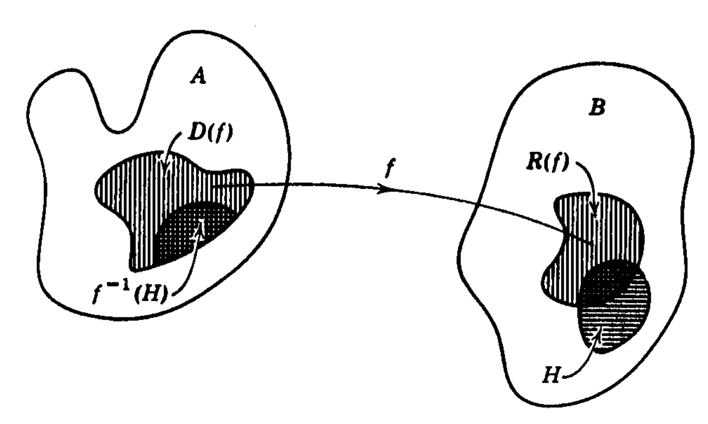


Figura 2.7. Imágenes inversas.

2.12 DEFINICION. Si H es un subconjunto de B, entonces, la **imagen inversa** de H bajo f es el subconjunto de $D(f)$ dado por

$$\{x : f(x) \in H\}.$$

Normalmente, la imagen inversa de un conjunto H bajo f se denota mediante el simbolo $f^{-1}(H)$.(véase la Fig. 2.7.)

De nuevo se hace hincapié en que f puede no ser inyectiva, de manera que la función inversa f^{-1} puede no existir.(Sin embargo, si f^{-1} existe, entonces f^{-1} (H) es la imagen directa de H bajo f^{-1}).

2.13 TEOREMA. *Sea f una función con dominio en A y rango en B y sean G, H subconjuntos de B.*

(a) Si $G \subseteq H$, entonces $f^{-1}(G) \subseteq f^{1}(H)$. (b) $f^{-1}(g \cap H) = f^{-1}(G) \cap f^{-1}(H)$.
(c) $f^{-1}(G \cup H) = f^{-1}(G) \cup f^{-1}(H)$. (d) $f^{-1}(G \setminus H) = f^{-1}(G) \setminus f^{-1}(H)$.

DEMOSTRACION. *(a)* Suponga que $x \in f^{-1}$ *(G);* entonces $f(x) \in G \subseteq H$, de aquí que $x \in f^{-1}(H)$.

(b) Puesto que $G \cap H$ es un subconjunto de G y de H, de la parte *(a)* se infiere que

$$f^{-1}(G \cap H) \subseteq f^{-1}(G) \cap f^{-1}(H).$$

Inversamente, si $x \in f^{-1}(G) \cap f^{-1}(H)$, entonces $f(x) \in G$ y $f(x) \in H$. Por lo tanto, $f(x) \in G \cap H$ y $x \in f^{-1}(G \cap H)$.

(c) Puesto que G y H son subconjuntos de $G \cup H$, de la parte *(a)* se deduce

$$f^{-1}(G \cup H) \supseteq f^{-1}(G) \cup f^{-1}(H).$$

Inversamente, si $x \in f^{-1}(G \cup H)$, entonces $f(x) \in G \cup H$. Se deduce que

$f(x)\in G$, de ahí $x\in f^{-1}(G)$, o bien $f(x)\in H$ en cuyo caso $x\in f^{-1}(H)$. Por lo tanto,

$$f^{-1}(G\cup H)\subseteq f^{-1}(G)\cup f^{-1}(H).$$

(d) Si $x\in f^{-1}(G\setminus H)$, entonces $f(x)\in G\setminus H$. Por lo tanto, $x\in f^{-1}(G)$ y $x\notin f^{-1}(H)$, de donde se infiere que

$$f^{-1}(G\setminus H)\subseteq f^{-1}(G)\setminus f^{-1}(H).$$

Inversamente si $w\in f^{-1}(G)\setminus f^{-1}(H)$, entonces $f(w)\in G$ y $f(w)\notin H$. De donde $f(w)\in G\setminus H$ y se deduce que

$$f^{-1}(G)\setminus f^{-1}(H)\subseteq f^{-1}(G\setminus H).$$

Q.E.D.

Ejercicios

2.A. Demuestre que la definición 2.2 de hecho da una función y no solamente un subconjunto.

2.B. Sea $A=B=\boldsymbol{R}$ y considere al subconjunto $C=\{(x,y):x^2+y^2=1\}$ de $A\times B$. ¿Es este conjunto una función con dominio en $\boldsymbol{R}$ y rango en $\boldsymbol{R}$?

2.C. Considere al subconjunto de $\boldsymbol{R}\times\boldsymbol{R}$ definido por $D=\{(x,y):|x|+|y|=1\}$. Describa este conjunto con palabras. ¿Es éste una función?

2.D. Dé un ejemplo de dos funciones f, g de $\boldsymbol{R}$ a $\boldsymbol{R}$ tales que $f\neq g$, pero $f\circ g=g\circ f$.

2.E. Demuestre que si f es una inyección de A a B, entonces $f^{-1}=\{(b,a):(a,b)\in f\}$ es una función. Después demuestre que es una inyección.

2.F. Suponga que f es una inyección. Demuestre que $f^{-1}\circ f(x)=x$ para toda x en $D(f)$ y $f\circ f^{-1}(y)=y$ para toda y en $R(f)$.

2.G. Sean f y g funciones y suponga que $g\circ f(x)=x$ para toda x en $D(f)$. Demuestre que f es inyección y que $R(f)\subseteq D(g)$ y $D(f)\subseteq R(g)$.

2.H. Sean f,g funciones tales que

$$g\circ f(x)=x \text{ para toda } x \text{ en } D(f),$$
$$f\circ g(y)=y \text{ para toda } y \text{ en } D(g).$$

Demuestre que $g=f^{-1}$.

2.I. Demuestre que la imagen directa $f(E)$ es vacía si y sólo si $E\cap D(f)=\emptyset$.

2.J. Sea f la función de $\boldsymbol{R}$ en $\boldsymbol{R}$ dada por $f(x)=x^2$ y sean $E=\{x\in\boldsymbol{R}:-1\le x\le 0\}$ y $F=\{x\in\boldsymbol{R}:0\le x\le 1\}$. Entonces $E\cap F=\{0\}$ y $f(E\cap F)=$ mientras que $f(E)=f(F)=\{y\in\boldsymbol{R}:0\le y\le 1\}$. Por lo tanto, $f(E\cap F)$ es un subconjunto propio de $f(E)\cap f(F)$. Ahora suprima el 0 de E y F.

2.K. Si f, E, F son como en el ejercicio 2.J, entonces $E\setminus F=\{x\in\boldsymbol{R}:-1\le x<0\}$ y $f(E)\setminus f(F)=\emptyset$. Por lo tanto, *no* se deduce que

$$f(E\setminus F)\subseteq f(E)\setminus f(F).$$

2.L. Demuestre que si f es una inyección de $D(f)$ en $R(f)$ y si H es un subconjunto

de $R(f)$, entonces la imagen inversa de H bajo f coincide con la imagen directa de H bajo la función inversa f^{-1}.

2.M. Si f y g son como en la definición 2.2, entonces $D(g \circ f) = f^{-1}(D(g))$.

Sección 3 Conjuntos finitos e infinitos

El propósito de esta sección es muy limitado: se verán los términos "finito", "contable" e "infinito". Se proporciona una base para el estudio de los números cardinales, pero no se prosigue dicho análisis. A pesar de que las teorías de números cardinales y ordinales son fascinantes en sí mismas, sólo un ligero conocimiento de estos temas resulta verdaderamente indispensable para comprender el material de este libro.†

Se supone que el lector está familiarizado con el conjunto de *números naturales*. Se denota a este conjunto mediante el símbolo N; los elementos de N se designan con los conocidos símbolos

$$1, 2, 3, \ldots.$$

Si n $m \in N$, todos tenemos una idea intuitiva de lo que significa la afirmación n es menor que o igual a m. Se usará esta afirmación teniendo en cuenta que una precisión total requiere más análisis del que se ha dado. Se supondrá que *todo subconjunto no vacío de N tiene un elemento mínimo.* Esta es una propiedad importante de N; en algunas ocasiones se dice que N es un subconjunto *bien ordenado* para significar que N tiene esta propiedad. Esta propiedad del buen orden es equivalente a la *inducción matemática.* Se emplearán argumentos basados en inducción matemática que supone le son familiares al lector.

Un **segmento inicial** de N es un conjunto que consta de todos los números naturales que son menores o iguales a algún elemento fijo de N. De este modo, un segmento inicial S_n de N determina y es determinado por un elemento n de N de la siguiente manera:

Un elemento x de N pertenece a S_n si y sólo si $x \leq n$.

Por ejemplo: el subconjunto $S_2 = \{1, 2\}$ es el segmento inicial de N determinado por el número natural 2; el subconjunto $S_4 = \{1, 2, 3, 4\}$ es el segmento inicial de N determinado por el número natural 4; pero el subconjunto $\{1, 3, 5\}$ de N *no* es un segmento inicial de N ya que contiene a 3 pero no a 2, y a 5 pero no a 4.

3.1 DEFINICION Un conjunto B es **finito** si es vacío o si hay una biyección con dominio B y rango en un segmento inicial de N. Si no existe tal función, el conjunto es **infinito.** Si hay una biyección de B sobre N, entonces el

† Véanse los libros de Halmos y Hamilton - Landin citados en la bibliografía.

conjunto B es **numerable** (o **enumerable**). Si un conjunto es finito o es numerable, se dice que es **contable.**

Cuando hay una función inyectiva (uno a uno) con dominio B y rango C, a veces se dice que B se puede poner en *correspondencia uno a uno* con C. Usando esta terminología, se reformula la definición 3.1 y se dice que un conjunto B es finito si es vacío o si se puede poner en correspondencia uno a uno con un subconjunto de un segmento inicial de N. Se dice que B es numerable si se puede poner en correspondencia uno a uno con todo N.

Se observará que, por definición, un conjunto B es finito o bien infinito. Sin embargo, puede ser que debido a la descripción del conjunto no resulte trivial decidir si el conjunto dado B es finito o infinito.

Los subconjuntos de N que se denotan por $\{1,3,5\}, \{2,4,5,8,10\}, \{2, 3, \ldots, 100\}$, son finitos ya que, a pesar de que no son segmentos iniciales de N, están contenidos en segmentos iniciales de N y de ahí que se puedan poner en correspondencia uno a uno con subconjuntos de segmentos iniciales de N. El conjunto E de números naturales pares

$$E = \{2, 4, 6, 8, \ldots\}$$

y el conjunto O de números naturales impares

$$O = \{1, 3, 5, 7, \ldots\}$$

no son segmentos iniciales de N, Sin embargo, dado que se pueden poner en correspondencia uno a uno con todo N (¿cómo?), ambos son numerables.

A pesar de que el conjunto $\mathbf{Z}$ de todos los enteros

$$\mathbf{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\},$$

contiene al conjunto N, se puede ver que $\mathbf{Z}$ es un conjunto numerable. (¿cómo?)

Ahora se enunciarán algunos teoremas sin demostración. Al leerlos por primera vez, probablemente a lo mejor sea aceptarlos sin analizarlos en detalle; sin embargo, en una revisión posterior el lector hará bien en intentar proporcionar demostraciones para estas aseveraciones. Al hacer esto, encontrará útil la propiedad inductiva del conjunto N de números naturales. †

3.2 TEOREMA. *Un conjunto B es contable si y sólo si hay una inyección con dominio B y rango en N.*

3.3 TEOREMA. *Cualquier subconjunto de un conjunto finito es finito. Cualquier subconjunto de un conjunto contable es contable.*

3.4 TEOREMA. *La unión de una colección finita de conjuntos finitos es un conjunto finito. La unión de una colección contable de conjuntos contables es un conjunto contable.*

†El lector que desee aprender acerca de estos temas deberá consultar el libro de Halmos citado en la bibliografía.

Como resultado de la segunda parte del teorema 3.4 se tiene que el conjunto $\boldsymbol{Q}$ de todos los números racionales forma un conjunto contable. (Recuerde que un número racional es una fracción m/n, en donde m y n son enteros y $n \neq 0$.) Para ver que $\boldsymbol{Q}$ es un conjunto contable se forman los conjuntos:

$$A_0 = \{0\},$$
$$A_1 = \{\tfrac{1}{1}, -\tfrac{1}{1}, \tfrac{2}{1}, -\tfrac{2}{1}, \tfrac{3}{1}, -\tfrac{3}{1}, \ldots\},$$
$$A_2 = \{\tfrac{1}{2}, -\tfrac{1}{2}, \tfrac{2}{2}, -\tfrac{2}{2}, \tfrac{3}{2}, -\tfrac{3}{2}, \ldots\},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$A_n = \left\{\frac{1}{n}, -\frac{1}{n}, \frac{2}{n}, -\frac{2}{n}, \frac{3}{n}, -\frac{3}{n}, \ldots\right\},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

Observe que cada uno de los conjuntos A_n es contable y que la unión de éstos es todo $\boldsymbol{Q}$. Por lo tanto, el teorema 3.4 afirma que $\boldsymbol{Q}$ es contable. De hecho, se puede enumerar $\boldsymbol{Q}$ por medio del "procedimiento de la diagonal":

$$0, \tfrac{1}{1}, -\tfrac{1}{1}, \tfrac{1}{2}, \tfrac{2}{1}, -\tfrac{1}{2}, \tfrac{1}{3}, \ldots .$$

Usando este tipo de argumento, el lector deberá ser capaz de elaborar una demostración para el teorema 3.4. Vea también el ejercicio 3.K.

La incontabilidad de R y de I

A pesar de que el conjunto de números racionales es contable, la totalidad del conjunto $\boldsymbol{R}$ de números reales no es contable. De hecho, el conjunto $\boldsymbol{I}$ de números reales x que satisfacen $0 \leq x \leq 1$ no es contable. Para probar esto se utilizará la elegante demostración de la "diagonal" de G. Cantor.† Se supone que se sabe que todo número real x tal que $0 \leq x \leq 1$ tiene una representación decimal de la forma $x = 0.a_1a_2a_3 \cdots$, en donde cada a_k designa uno de los dígitos 0,1,2,3,4,5,6,7,8,9. Se debe observar que ciertos números reales tienen dos representaciones de esta forma (por ejemplo, el número racional 1/10 tiene las dos representaciones

$$0.1000 \cdots \quad \text{y} \quad 0.0999 \cdots).$$

Se podría elegir alguna de estas dos representaciones, pero no es necesario hacerlo. Dado que hay una infinidad de números racionales en el intervalo $0 \leq x \leq 1$, (¿por qué?), el conjunto $\boldsymbol{I}$ no puede ser finito. Ahora se demostrará que no es numerable. Suponga que hay una enumeración $x_1, x_2, x_3, \ldots$ de todos los números reales que satisfacen $0 \leq x \leq 1$ dados por

† GEORG CANTOR (1845-1918) nació en San Petersburgo, estudió en Berlín con Weierstrass y ejerció el magisterio en Halle. Es conocido sobre todo por su trabajo en teoría de conjuntos que llevó a cabo durante los años 1874-1895.

$$x_1 = 0.a_1a_2a_3 \cdots$$
$$x_2 = 0.b_1b_2b_3 \cdots$$
$$x_3 = 0.c_1c_2c_3 \cdots$$
$$\cdots\cdots\cdots\cdots$$

Ahora bien, sea y_1 un dígito distinto de 0, a_1, y 9; sea y_2 un dígito distinto de 0, b_2, y 9; sea y_3 un dígito distinto de 0, c_3, y 9, etc. Considere el número y con representación decimal

$$y = 0.y_1y_2y_3 \cdots$$

que claramente satisface $0 \leq y \leq 1$. El número y no es uno de los números con doble representación decimal, ya que $y_n \neq 0$, 9. Además $y \neq x_n$ para cualquier n (ya que los n-ésimos dígitos en las representaciones decimales de y y x_n son diferentes). Por lo tanto, cualquier colección numerable de números reales en este intervalo omitirá cuando menos a un número real perteneciente a este intervalo. En consecuencia, este intervalo no es un conjunto contable.

Suponga que un conjunto A es infinito, se dará *por establecido* que hay una correspondencia uno a uno entre un subconjunto A y todo N. En otras palabras, *se da por supuesto que todo conjunto infinito contiene un subconjunto numerable*. Esta afirmación es un modelo débil del llamado "axioma de selección". que es uno de los axiomas comunes de la teoría de conjuntos. Una vez que el lector haya asimilado el contenido de este libro, podrá proceder al estudio analítico de los fundamentos que se han estado dando de manera un tanto informal. Sin embargo, por el momento hará bien en considerar esta última afirmación como un axioma temporal. Más adelante se podrá reemplazar por un axioma más ambicioso de la teoría de conjuntos.

Ejercicios

3.A. Muestre una correspondencia uno a uno entre el conjunto E de números naturales pares y N.

3.B. Muestre una correspondencia uno a uno entre el conjunto O de números naturales impares y N.

3.C. Muestre una correspondencia uno a uno entre N y un subconjunto propio de N.

3.D. Si A está contenido en algún segmento inicial de N, utilice la propiedad del buen orden de N para definir una biyección de A sobre algún segmento inicial de N.

3.E. Dé un ejemplo de una colección contable de conjuntos finitos cuya unión no sea finita.

3.F. Utilice el hecho de que todo conjunto infinito tiene un subconjunto numerable para demostrar que todo conjunto infinito se puede poner en correspondencia uno a uno con un subconjunto propio de sí mismo .

3.G. Demuestre que si el conjunto A se puede poner en correspondencia uno a uno con un conjunto B, entonces B se puede poner en correspondencia uno a uno con

3.H. Demuestre que si el conjunto A se puede poner en correspondencia uno a uno con un conjunto B, y si B se puede poner en correspondencia uno a uno con un conjunto C, entonces A se puede poner en correspondencia uno a uno con C.

3.I. Usando inducción en $n \in N$, demuestre que el segmento inicial determinado por n no se puede poner en correspondencia uno a uno con el segmento inicial determinado por $m \in \mathbf{N}$, si $m < n$.

3.J. Demuestre que N no se puede poner en correspondencia uno a uno con ningún segmento inicial de N.

3.K. Para cada $n \in \mathbf{N}$ sea $A_n = \{a_{nj} : j \in \mathbf{N}\}$, y suponga que $A_n \cap A_m = \emptyset$ para $n \neq m$, $n, m \in \mathbf{N}$. Demuestre que la función $f(n, j) = \frac{1}{2}(n + j - 2)(n + j - 1) + n$ da una enumeración de $\bigcup\{A_n : n \in \mathbf{N}\}$.

I
LOS NUMEROS REALES

En este capítulo se analizarán las propiedades del sistema de números reales. Aunque sería posible construir este sistema a partir de un conjunto más elemental (como el conjunto $\boldsymbol{N}$ de números naturales o el conjunto $\boldsymbol{Q}$ de números racionales), no se hará así. En vez de ello, se dará una lista de propiedades relacionadas con el sistema de números reales y se mostrará cómo es posible deducir otras propiedades a partir de aquellas que se han establecido.

Para mayor claridad, es preferible no dar todas las propiedades del sistema de números reales simultáneamente. En lugar de ello, se introducen primero, en la sección 4, las "propiedades algebraicas" basadas en las dos operaciones de la suma y de la multiplicación y se analizan en forma breve algunas de sus consecuencias. En seguida se presentan las "propiedades de orden". En la sección 6 se da el último paso agregando la "propiedad de completación". Existen varias razones por las que se usa este procedimiento un tanto seccionado. En primer lugar, son varias propiedades las que se deben tomar en cuenta y es conveniente escoger sólo unas cuantas a la vez. Además, las demostraciones que se requieren en los pasos algebraicos preliminares al principio son más naturales que algunas de las demostraciones posteriores. Por último, puesto que existen algunos otros métodos interesantes para agregar la "propiedad de completación" se desea mantenerla aislada del resto de los supuestos.

En parte, el propósito de las secciones 4 y 5 es proporcionar ejemplos de demostraciones de teoremas elementales que se derivan de supuestos explícitamente establecidos. Por experiencia se sabe que los estudiantes que no han experimentado mucho con demostraciones rigurosas pueden asimilar fácilmente los argumentos que se presentan en estas secciones y después entrar a la sección 6. *Por otro lado, los alumnos que estén familiarizados con el método axiomático y la técnica de las demostraciones podrán pasar a la sección 6 después de un repaso rápido a las secciones 4 y 5.*

En la sección 7 se introduce el concepto de una cortadura en el sistema de números reales y se definen varios tipos de celdas e intervalos. Se establece la importante propiedad de celdas nidificadas de $\boldsymbol{R}$ y se analiza brevemente el conjunto de Cantor.

Sección 4 Las propiedades algebraicas de R

En esta sección se dará la estructura "algebraica" del sistema de números reales. Resumiendo, se puede decir que los números reales forman un "campo" en el sentido del álgebra abstracta. En seguida se explica lo que esto significa.

Una **operación binaria** en un conjunto F se entiende que es una función B con dominio $F \times F$ y rango en F. En vez de usar la notación $B(a, b)$ para denotar el valor de la operación binaria B en el punto (a, b) en $F \times F$, es convencional usar una notación tal como aBb, o $a + b$, o $a \cdot b$.

4.1 PROPIEDADES ALGEBRAICAS DE R. En el conjunto R de números reales hay dos operaciones binarias (designadas por $+$ y $\cdot$ que se llaman suma y multiplicación respectivamente, que satisfacen las siguientes[1] propiedades:

(A1) $a + b = b + a$ para todas a, b en R;
(A2) $(a + b) + c = a + (b + c)$ para todas a, b, c en R;
(A3) existe un elemento O en R tal que $O + a = a$ y $a + O = a$ para toda a en R;
(A4) para cada elemento a en R hay un elemento $-a$ en R tal que $a + (-a) = O$ y $(-a) + a = 0$;
(M1) $a \cdot b = b \cdot a$ para todas a, b en R;
(M2) $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para todas a,b,c en R;
(M3) el elemento 1 en R es distinto de 0 y tiene la propiedad de que $1 \cdot a = a$ y $a \cdot 1 = a$ para toda a en R;
(M4) para cada elemento $a \neq 0$ en R hay un elemento $1/a$ en R tal que $a \cdot (1/a) = 1$ y $(1/a) \cdot a = 1$;
(D) $a \cdot (b + c) = (a \cdot b) + (a \cdot c)$ y $(b + c) \cdot a = (b \cdot a) + (c \cdot a)$ para todas a,b,c en R.

Estas propiedades seguramente le son familiares al lector. Ahora se obtendrán algunas consecuencias sencillas (pero importantes) de dichas propiedades. En primer lugar se demostrará que 0 es el único elemento de R que satisface *(A3)* y 1 es el único elemento que satisface *(M3)*.

4.2 TEOREMA. *(a) Si z y a son elementos de R tales que $z + a = a$, entonces $z = 0$.*

(b) Si w y $b \neq 0$ son elementos de R tales que $w \cdot b = b$, entonces $w = 1$.

DEMOSTRACION. *(a)* La hipótesis es que $z + a = a$. Sumar $-a$ a ambos lados y emplear *(A4)*, *(A2)*, *(A4)* y *(A3)* para obtener

*No se pretende que esta lista sea "mínima". Así las segundas afirmaciones en (A3) y (A4) se deducen de las primeras usando (A1).

$$0 = a + (-a) = (z + a) + (-a) = z + (a + (-a))$$
$$= z + 0 = z.$$

La demostración de la parte *(b)* se deja como ejercicio. Observe que se usa la hipótesis de que $b \neq 0$. Q.E.D.

En seguida se demuestra que los elementos $-a$ y $1/a$ (cuando $a \neq 0$) están determinados de manera única por las propiedades dadas en *(A4)* y *(M4)*.

4.3 TEOREMA. *(a) Si a y b son elementos de* **R** *y $a + b = 0$, entonces $b = -a$.*
(b) Si $a \neq 0$ y b son elementos de ***R*** *y $a \cdot b = 1$, entonces $b = 1/a$.*

DEMOSTRACION. *(a)* Si $a + b = 0$, sumar $-a$ a ambos lados para obtener $(-a) + (a + b) = -a + 0$. Ahora, usar *(A2)* en el lado izquierdo y *(A3)* en el derecho para obtener

$$((-a) + a) + b = -a.$$

Si se usan *(A4)* y *(A3)* en el lado izquierdo, se obtiene $b = -a$.

La demostración de *(b)* se deja como ejercicio. Observe que se usa la hipótesis de que $a \neq 0$. Q.E.D.

Las propiedades *(A4)* y *(M4)* garantizan la posibilidad de resolver las ecuaciones

$$a + x = 0, \qquad a \cdot x = 1 \quad (a \neq 0),$$

para x, y el teorema 4.3 implica la unicidad de las soluciones. Ahora se demuestra que los lados derechos de estas ecuaciones pueden ser elementos arbitrarios de ***R***.

4.4 TEOREMA. *(a) Sean a,b elementos arbitrarios de* ***R***. *Entonces la ecuación $a + x = b$ tiene como solución única $x = (-a) + b$.*
(b) Sean $a \neq 0$ y b elementos arbitrarios de ***R***. *Entonces la ecuación $a \cdot x = b$ tiene la solución única $x = (1/a) \cdot b$.*

DEMOSTRACION. Puesto que $a + ((-a) + b) = (a + (-a)) + b = 0 + b = b$, es claro que $x = (-a) + b$ es una solución de la ecuación $a + x = b$. Para ratificar que es la única solución, sea x_1 cualquier solución de esta ecuación; entonces

$$a + x_1 = b.$$

Se suma $-a$ a ambos lados para obtener

$$(-a) + (a + x_1) = (-a) + b.$$

Si se utilizan *(A3)*, *(A4)* y *(A2)*, se obtiene

$$x_1 = 0 + x_1 = (-a + a) + x_1$$
$$= (-a) + (a + x_1) = (-a) + b.$$

De donde $x_1 = (-a) + b$.

La demostración de la parte *(b)* se deja como ejercicio. Q.E.D.

4.5 TEOREMA. *Si a y b son cualesquier elementos de* $\mathbf{R}$, *entonces*

(a) $a \cdot 0 = 0$; (b) $-a = (-1) \cdot a$;
(c) $-(a+b) = (-a) + (-b)$; (d) $-(-a) = a$;
(e) $(-1) \cdot (-1) = 1$.

DEMOSTRACION. *(a)* A partir de *(M3)* se sabe que $a \cdot 1 = a$. De donde

$$a + a \cdot 0 = a \cdot 1 + a \cdot 0 = a \cdot (1 + 0)$$
$$= a \cdot 1 = a.$$

Si se aplica el teorema 4.2*(a)*, se deduce que $a \cdot 0 = 0$.

(b) Está visto que

$$a + (-1) \cdot a = 1 \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a$$
$$= 0 \cdot a = 0.$$

Del teorema 4.3*(a)* se deduce que $(-1) \cdot a = -a$.

(c) Se tiene

$$-(a+b) = (-1) \cdot (a+b) = (-1) \cdot a + (-1) \cdot b$$
$$= (-a) + (-b).$$

(d) De *(A4)* se tiene $(-a) + a = 0$. Conforme a la afirmación de unicidad del teorema 4.3*(a)*, se infiere que $a = -(-a)$.

(e) En la parte *(b)*, substituir $a = -1$. Se tiene

$$-(-1) = (-1) \cdot (-1).$$

Por lo tanto, la afirmación se deduce de la parte *(d)* con $a = 1$. Q.E.D.

4.6 TEOREMA *(a) Si $a \in \mathbf{R}$ y $a \neq 0$, entonces* $1/a \neq 0$ *y* $1/(1/a) = a$.

(b) Si a, $b \in \mathbf{R}$ y $a \cdot b = 0$, entonces $a = 0$ o bien $b = 0$.

(c) Si a, $b \in \mathbf{R}$, entonces $(-a) \cdot (-b) = a \cdot b$

(d) Si $a \in \mathbf{R}$ y $a \neq 0$, entonces $1/(-a) = -(1/a)$.

DEMOSTRACION. *(A)* Si $a \neq 0$, entonces $1/a \neq 0$ pues de no ser así $1 = a \cdot (1/a) = a \cdot 0 = 0$, contrario a *(M3)*. Dado que $(1/a) \cdot a = 1$, del teorema 4.3 *(b)* se infiere que $a = 1/(1/a)$.

(b) Suponga que $a \cdot b = 0$ y que $a \neq 0$. Si se multiplica por $1/a$, se obtiene

$$b = 1 \cdot b = ((1/a) \cdot a) \cdot b = (1/a) \cdot (a \cdot b)$$
$$= (1/a) \cdot 0 = 0.$$

Una demostración análoga es válida si $b \neq 0$.

(c) Por el teorema 4.5, se tiene $-a = (-1) \cdot a$, y $-b = (-1) \cdot b$; de aquí que

$$(-a) \cdot (-b) = ((-1) \cdot a) \cdot ((-1) \cdot b)$$
$$= (a \cdot (-1)) \cdot ((-1) \cdot b)$$
$$= a \cdot ((-1) \cdot (-1)) \cdot b = a \cdot 1 \cdot b$$
$$= a \cdot b.$$

(d) Si $a \neq 0$, entonces $1/a \neq 0$ y $-a \neq 0$. Como $a \cdot (1/a) = 1$, de la parte *(c)* se deduce que $(-a) \cdot (-(1/a)) = 1$. Aplicando el teorema 4.3 *(b)* se deduce que $1/(-a) = -(1/a)$ siendo esto lo que se afirma. Q.E.D.

Numeros racionales

De ahora en adelante por lo general se habrá de omitir el uso del punto para designar la multiplicación y se escribirá ab en vez de $a \cdot b$. Como de costumbre se escribirá a^2 para designar aa, a^3 para designar $aaa = (a^2)a$, y si $n \in \boldsymbol{N}$ se define $a^{n+1} = (a^n)a$. Usando inducción matemática se deduce que si $m, n \in \boldsymbol{N}$, entonces

(*) $$a^{m+n} = a^m a^n$$

para cualquier $a \in \boldsymbol{R}$. Análogamente, se escribirá 2 para designar $1+1$, 3 para designar $2+1 = (1+1)+1$, y así sucesivamente. Además, por lo general se escribirá $b-a$ en vez de $(-a)+b = b+(-a)$ y si $a \neq 0$, por lo regular se escribirá

$$b/a \quad \text{ó} \quad \frac{b}{a}$$

en vez de $(1/a) \cdot b = b \cdot (1/a)$. También se escribirá a^{-1} para designar $1/a$, y a^{-n} para designar $1/a^n$. Entonces se podrá demostrar que la fórmula anterior es válida para $m, n \in \boldsymbol{Z}$ cuando $a \neq 0$.

Los elementos de $\boldsymbol{R}$ que son de la forma

$$\frac{b}{a} \quad \text{ó} \quad \frac{-b}{a}$$

para $a, b \in \boldsymbol{N}$, $a \neq 0$, se dice que son **números racionales**, y el conjunto de todos los números racionales en $\boldsymbol{R}$ se designará con la notación clásica $\boldsymbol{Q}$. To-

dos los elementos de $\boldsymbol{R}$ que *no* son números racionales se dice que son **números irracionales.** A pesar de que esta terminología no es muy apropiada, es bastante clásica y se habrá de adoptar.

Se concluirá esta sección con una demostración del hecho de que no existe un número racional cuyo cuadrado sea 2.

√ 4.7 TEOREMA. *No existe ningún número racional r tal que* $r^2 = 2$.

DEMOSTRACION. Suponga, por el contrario, que $(p/q)^2 = 2$, en donde p y q son enteros. Sin pérdida de generalidad se puede suponer que p y q no tienen factores integrales en común. (¿por qué?) Dado que $p^2 = 2q^2$, se deduce que p debe ser un entero par (ya que si $p = 2k + 1$ es impar, entonces $p^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ es impar). Por lo tanto $p = 2k$ para algún entero k y en consecuencia $4k^2 = 2q^2$. Se deduce que $q^2 = 2k^2$, por lo que q también debe ser par. Por lo tanto, p y q son divisibles entre 2, lo que contradice a la hipótesis. Q.E.D.

Ejercicios.

4.A. Demostrar la parte *(b)* del teorema 4.2.
4.B. Demostrar la parte *(b)* del teorema 4.3.
4.C. Demostrar la parte *(c)* del teorema 4.4.

4.D. Usando inducción matemática, demostrar que si $a \in \boldsymbol{R}$ y m, $n \in \boldsymbol{N}$, entonces $a^{m+n} = a^m a^n$.

4.E. Demostrar que si $a \in \boldsymbol{R}$, $a \neq 0$, y $m, n \in \boldsymbol{Z}$, entonces $a^{m+n} = a^m a^n$.

4.F. Emplear el argumento del teorema 4.7 para demostrar que no existe un número racional s tal que $s^2 = 6$.

✓ 4.G. Modificar el argumento del teorema 4.7 para demostrar que no existe un número racional t tal que $t^2 = 3$.

✓ 4.H. Si $\xi \in \boldsymbol{R}$ es irracional y $r \in \boldsymbol{R}$, $r \neq 0$, es racional, demostrar que $r + \xi$ y $r\xi$ son irracionales.

Sección 5 Las propiedades de orden de $\boldsymbol{R}$

El propósito de esta sección es introducir las propiedades de "orden" de $\boldsymbol{R}$, las cuales desempeñarán un papel muy importante en las secciones siguientes. La manera más sencilla de introducir el concepto de orden es haciendo uso del concepto de "positividad estricta" que en seguida se explica.

5.1 LAS PROPIEDADES DE ORDEN DE $\boldsymbol{R}$. Existe un subconjunto no vacío P de $\boldsymbol{R}$ llamado el conjunto de **números reales estrictamente positivos** que satisface las siguientes propiedades:

(i) Si a,b pertenecen a P, entonces $a+b$ pertenece a P.

(ii) Si a,b pertenecen a P, entonces ab pertenece a P.

(iii) Si a pertenece a $\boldsymbol{R}$, entonces una de las siguientes relaciones se cumple exactamente: $a \in P$, $a = 0$, $-a \in P$.

A la condición *(iii)* algunas veces se la llama **propiedad de tricotomía.** Esta implica que el conjunto $N = \{-a : a \in P\}$, algunas veces llamado conjunto de **números reales estrictamente negativos,** no tiene elementos en común con P. De hecho, todo el conjunto $\boldsymbol{R}$ es la unión de los tres conjuntos ajenos P, $\{0\}$, N.

5.2 DEFINICION. Si $a \in P$, se dice que a es **un número real estrictamente positivo** y se escribe $a > 0$. Si a está en P o bien es 0, se dice que a es un **número real positivo** y se escribe $a \geq 0$. Si $-a \in P$, se dice que a es un **número real estrictamente negativo** y se escribe $a < 0$. Si $-a$ está en P o bien es 0, se dice que a es un **número real negativo** y se escribe $a \leq 0$.

Se debe observar que, de acuerdo con la terminología que se acaba de implantar, el número 0 es tanto positivo como negativo; es el único número con dicha categoría dual. Esta terminología puede parecer al principio un poco extraña, pero resultará conveniente. Algunos autores reservan el término "positivo" para los elementos del conjunto P y utilizan el término "no negativo" para los elementos de $P \cup \{0\}$.

En seguida se darán las relaciones de orden.

5.3 DEFINICION. Sean a,b elementos de $\boldsymbol{R}$. Si $a - b \in P$, entonces se escribe $a > b$. Si $-(a-b) \in P$, entonces se escribe $a < b$. Si $a - b \in P \cup \{0\}$, entonces se escribe $a \geq b$. Si $-(a-b) \in P \cup \{0\}$, entonces se escribe $a \leq b$.

Como es costumbre, es conveniente invertir los signos y escribir

$$b < a, \qquad b > a, \qquad b \leq a, \qquad b \geq a,$$

respectivamente. Además, si $a < b$ y $b < c$, entonces con frecuencia se escribe

$$a < b < c \qquad \text{ó} \qquad c > b > a.$$

Si $a \leq b$ y $b < c$, entonces con frecuencia se escribe

$$a \leq b < c \qquad \text{ó} \qquad c > b \geq a.$$

Propiedades de orden

Se establecerán ahora las propiedades básicas de la relación de orden en $\boldsymbol{R}$. Estas son las conocidas "leyes" para las desigualdades que el lector ha visto en cursos anteriores. Se usarán con frecuencia en secciones posteriores y son de gran importancia.

5.4 TEOREMA. *Sean a,b,c elementos de* $\mathbf{R}$.

(a) Si $a > b$ *y* $b > c$, *entonces* $a > c$.

(b) Exactamente una de las siguientes relaciones es válida: $a > b$, $a = b$, $a < b$.

(c) Si $a \geq b$ *y* $b \geq a$, *entonces* $a = b$.

DEMOSTRACION. *(a)* Si $a - b$ y $b - c$ pertenecen a P, entonces de 5.1 *(i)* se deduce que $a - c = (a - b) + (b - c)$ también pertenece a P. De donde $a > c$.

(b) Por 5.1 *(iii)* sucede exactamente una de las siguientes posibilidades: $a - b \in P$, $a - b = 0$, $b - a = -(a - b) \in P$.

(c) Si $a \neq b$, entonces, por el inciso *(b)*, se debe tener $a - b$ o bien $b - a$ en P. Por lo tanto, $a > b$ o bien $b > a$. En cualquiera de los dos casos una de las hipótesis se contradice. Q.E.D.

5.5 TEOREMA. *(a)* Si $0 \neq a \in \mathbf{R}$, entonces $a^2 > 0$.

(b) $1 > 0$.

(c) Si $n \in \mathbf{N}$, entonces $n > 0$.

DEMOSTRACION. *(a)* Ya sea que a o a pertenezca a P. Si $a \in P$ entonces por 5.1 *(ii)* se tiene $a^2 = aa \in P$. Si $-a \in P$, entonces por el teorema 4.7*(c)* se tiene $a^2 = (-a)(-a) \in P$. Por lo tanto, en cualquiera de los casos, $a^2 \in P$.

(b) Dado que $1 = (1)^2$, la conclusión se deduce de *(a)*.

(c) Se usa inducción matemática. La validez de la afirmación con $n = 1$ es la parte *(b)*. Si la afirmación es cierta para el número natural k (es decir, suponiendo $k \in P$), entonces, puesto que $1 \in P$, se deduce de 5.1 *(i)* que $k + 1 \in P$. Por lo tanto, la afirmación es verdadera para todos los números naturales. Q.E.D.

Las siguientes propiedades posiblemente le son familiares al lector.

5.6 TEOREMA. *Sean a,b,c,d elementos de* $\mathbf{R}$.

(a) Si $a > b$, *entonces* $a + c > b + c$.

(b) Si $a > b$ *y* $c > d$, *entonces* $a + c > b + d$.

(c) Si $a > b$ *y* $c > 0$, *entonces* $ac > bc$.

(c') Si $a > b$ *y* $c < 0$, *entonces* $ac < bc$.

(d) Si $a > 0$, *entonces* $1/a > 0$.

(d') Si $a < 0$, *entonces* $1/a < 0$.

DEMOSTRACION. *(a)* Observe que $(a + c) - (b + c) = a - b$.

(b) Si $a - b$ y $c - d$ pertenecen a P, entonces por 5.1*(i)* se concluye que $(a + c) - (b + d) = (a - b) + (c - d)$ también pertenece a P.

(c) Si $a - b$ y c pertenecen a P, entonces por 5.1 *(ii)* $ac - bc = (a - b)c$ también pertenece a P.

(c') Si $a-b$ y $-c$ pertenecen a P, entonces por 5.1*(ii)* $bc-ac=(a-b)(-c)$ también pertenece a P.

(d) Si $a>0$, entonces por 5.1*(iii)* $a\neq 0$ para que el elemento $1/a$ exista. Si $1/a=0$, entonces $1=a(1/a)=0$, contradicción. Si $1/a<0$, entonces la parte *(c')* con $c=1/a$ implica que $1=a(1/a)<0$, contradiciéndose 5.5*(b)*. Por lo tanto, se debe tener $1/a>0$, ya que las otras dos posibilidades se han excluido.

(d') Esta parte se puede demostrar usando un argumento similar al de la parte *(d)*. De manera alternativa se puede recurrir al teorema 4.6*(d)* y usar la parte *(d)* directamente. Q.E.D.

5.7 TEOREMA. Si $a>b$, entonces $a>\frac{1}{2}(a+b)>b$.

DEMOSTRACION. Dado que $a>b$, se deduce del teorema 5.6*(a)*, con $c=a$, que $2a>a+b$ y del teorema 5.6*(a)*, con $c=b$ que $a+b>2b$. Por el teorema 5.5*(c)* se sabe que $2>0$ y a partir del teorema 5.6*(d)* que $\frac{1}{2}>0$. Aplicado el teorema 5.6*(c)* con $c=\frac{1}{2}$, se deduce que $a>\frac{1}{2}(a+b)$ y $\frac{1}{2}(a+b)>b$. Por lo tanto, $a>\frac{1}{2}(a+b)>b$, como se sostiene. Q.E.D.

El teorema que se acaba de demostrar (con $b=0$) implica que dado cualquier número a estrictamente positivo, existe otro número estrictamente menor y estrictamente positivo (a saber, $\frac{1}{2}a$). De modo que *no existe ningún número real mínimo estrictamente positivo.*

Ya se ha visto que si $a>0$ y $b>0$, entonces $ab>0$. También si $a<0$ y $b<0$, entonces $ab>0$. En seguida se demostrará que lo inverso es cierto.

5.8 TEOREMA. *Si $ab>0$, entonces se tiene $a>0$ y $b>0$, o bien* $a<0$ y $b<0$,

DEMOSTRACION. Si $ab>0$, entonces $a\neq 0$ y $b\neq 0$. (¿por qué?) Si $a>0$ entonces del teorema 5.6*(d)* se deduce que $1/a>0$ y de 5.6*(c)* que $b=((1/a)a)b=(1/a)(ab)>0$. Por otro lado, si $a<0$, entonces del teorema 5.6*(d')* se deduce que $1/a<0$ y de 5.6*(c')* que $b=((1/a)a)b=(1/a)(ab)<0$. Q.E.D.

5.9 COROLARIO. *Si $ab<0$, entonces, se tiene $a>0$ y $b<0$ o bien se tiene a $a<0$ y $b>0$.*

La demostración de esta afirmación se deja como ejercicio.

Valor absoluto

La propiedad de tricotomía 5.1*(iii)* asegura que si $a\neq 0$, entonces uno de los números a y $-a$ es estrictamente positivo. El valor absoluto de $a\neq 0$ se define como el número estrictamente positivo del par $\{a, -a\}$, y el valor absoluto de 0 se define como 0.

5.10 DEFINICION. Si $a \in \boldsymbol{R}$, el *valor absoluto* de a se denota mediante el símbolo $|a|$ y está definido por

$$|a| = a \quad \text{si} \quad a \geq 0,$$
$$= -a \quad \text{si} \quad a < 0.$$

De modo que el dominio de la función valor absoluto es todo $\boldsymbol{R}$, su rango es $P \cup \{0\}$, y aplica los elementos a, $-a$, en el mismo elemento.

5.11 TEOREMA. *(a)* $|a| = 0$ si y sólo si $a = 0$.
(b) $|-a| = |a|$ *para toda* $a \in \boldsymbol{R}$.
(c) $|ab| = |a|\,|b|$ *para todas* $a, b \in \boldsymbol{R}$.
(d) Si $c \geq 0$, *entonces* $|a| \leq c$ *si y sólo si* $-c \leq a \leq c$.
(e) $-|a| \leq a \leq |a|$ *para toda* $a \in \boldsymbol{R}$.

DEMOSTRACION. *(a)* Si $a = 0$, entonces por definición $|0| = 0$. Si $a \neq 0$, entonces también $-a \neq 0$, de manera que $|a| \neq 0$.

(b) Si $a = 0$, entonces $|0| = a = |-0|$. Si $a > 0$, entonces $|a| = a = |-a|$. Si $a < 0$, entonces $|a| = -a = |-a|$.

(c) Si $a > 0$ y $b > 0$, entonces $ab > 0$ de manera que $|ab| = ab = |a|\,|b|$. Si $a < 0$ y $b > 0$, entonces $ab < 0$ de tal modo que $|ab| = -(ab) = (-a)b = |a|\,|b|$. Los otros casos se tratan de manera similar.

(d) Si $|a| \leq c$, entonces $a \leq c$ así como $-a \leq c$. (¿por qué?) A partir de esto último y del teorema 5.6*(c')* se deduce que $-c \leq a$ de tal modo que $-c \leq a \leq c$. De manera inversa si esta relación se sumple entonces $a \leq c$ así como $-a \leq c$, de donde $|a| \leq c$.

(e) Utilizar la parte *(d)* con $c = |a| \geq 0$. Q.E.D.

El siguiente resultado se usará con mucha frecuencia en lo sucesivo. (Recuerde que $a \pm b$ significa tanto $a + b$ como $a - b$.)

5.12 LA DESIGUALDAD DEL TRIANGULO. Si a,b son cualesquier números reales, entonces

$$\big|\,|a| - |b|\,\big| \leq |a \pm b| \leq |a| + |b|.$$

DEMOSTRACION. De acuerdo con el teorema 5.11*(e)* se tiene $-|a| \leq a \leq |a|$ y $-|b| \leq \pm b \leq |b|$. Utilizando 5.6*(b)* se deduce que

$$-(|a| + |b|) = -|a| - |b| \leq a \pm b \leq |a| + |b|.$$

A partir del teorema 5.11*(d)* se infiere que $|a \pm b| \leq |a| + |b|$, demostrando la segunda parte de la desigualdad.

Dado que $|a| = |(a - b) + b| \leq |a - b| + |b|$ (¿por qué?), se deduce que $|a| - |b| \leq |a - b|$. Análogamente $|b| - |a| \leq |a - b|$. (¿por que?) Combinando estas dos desigualdades, se deduce que $\big|\,|a| - |b|\,\big| \leq |a - b|$, que es la primera parte de la desigualdad con el signo menos. Para obtener la desigualdad con el signo más, sustituir b por $-b$.

5.13 COROLARIO. *Si $a_1, a_2, \ldots, a_n$ son cualesquier n números reales, entonces*

$$|a_1 + a_2 + \cdots + a_n| \le |a_1| + |a_2| + \cdots + |a_n|.$$

DEMOSTRACION. Si $n = 2$, la conclusión es precisamente 5.12. Si $n > 2$, se usa inducción matemática y el hecho de que

$$\begin{aligned}|a_1 + a_2 + \cdots + a_k + a_{k+1}| &= |(a_1 + a_2 + \cdots + a_k) + a_{k+1}|\\ &\le |a_1 + a_2 + \cdots + a_k| + |a_{k+1}|.\end{aligned}$$ Q.E.D.

Ejercicios

5.A. Si $a, b \in \mathbf{R}$ y $a^2 + b^2 = 0$, demostrar que $a = b = 0$.

5.B. Si $n \in \mathbf{N}$, demostrar que $n^2 \ge n$ y por lo tanto $1/n^2 \le 1/n$.

5.C. Si $a > -1$, $a \in \mathbf{R}$, demostrar que $(1 + a)^n \ge 1 + na$ para toda $n \in \mathbf{N}$. Esta desigualdad se llama la **desigualdad de Bernoulli.**† (*Sugerencia:* usar inducción matemática.)

5.D. Si $c > 1$, $c \in \mathbf{R}$, demostrar que $c^n \ge c$ para toda $n \in \mathbf{N}$. (*Sugerencia:* $c = 1 + a$ con $a > 0$).

5.E. Si $c > 1$, $c \in \mathbf{R}$, demostrar que $c^m \ge c^n$ para $m \ge n$, $m, n \in \mathbf{N}$.

5.F. Suponga que $0 < c < 1$. Si $m \ge n$, $m, n \in \mathbf{N}$, demostrar que $0 < c^m \le c^n < 1$.

5.G. Demostrar que $n < 2^n$ para toda $n \in \mathbf{N}$. y por lo tanto que $1/2^n < 1/n$ para toda $n \in \mathbf{N}$.

5.H. Si a y b son números reales positivos y $n \in \mathbf{N}$, entonces $a^n < b^n$ si y sólo si $a < b$.

5.I. Demostrar que si $a \le x \le b$ y $a \le y \le b$, entonces $|x - y| \le b - a$. Interpretar esto geométricamente.

5.J. Sea $\delta > 0$, $a \in \mathbf{R}$. Demostrar que $a - \delta < x < a + \delta$ si y sólo si $|x - a| < \delta$. Análogamente, $a - \delta \le x \le a + \delta$ si y sólo si $|x - a| \le \delta$.

5.K. Si $a, b \in \mathbf{R}$ y $b \ne 0$, demostrar que $|a/b| = |a|/|b|$.

5.L. Demostrar que si $a, b \in \mathbf{R}$, entonces $|a + b| = |a| + |b|$ si y sólo si $ab \ge 0$.

5.M. Graficar los puntos (x,y) en el plano $\mathbf{R} \times \mathbf{R}$ para los cuales $|y| = |x|$.

5.N. Graficar los puntos (x,y) en el plano $\mathbf{R} \times \mathbf{R}$ para los cuales $|x| + |y| = 1$.

5.O. Si x,y,z pertenecen a $\mathbf{R}$, entonces $x \le y \le z$ si y sólo si $|x - y| + |y - z| = |x - z|$.

5.P. Si $0 < a < 1$, entonces $0 < a^2 < a < 1$, mientras que si $1 < a$, entonces $1 < a < a^2$.

† JACOB BERNOULLI (1654-1705) fue miembro de una familia suiza que dio varios matemáticos que desempeñaron un papel importante en el desarrollo del cálculo.

Sección 6 La propiedad de complementación de R

En esta sección se presentará otra propiedad del sistema de números reales que a menudo se llama "propiedad de completación" ya que asegura la existencia de elementos en $\boldsymbol{R}$ cuando se satisfacen ciertas hipótesis. Existen diversas versiones de esta propiedad de completación. En este caso se ha escogido dar lo que tal vez sea el método más eficaz al suponer que conjuntos acotados en $\boldsymbol{R}$ tienen un supremo.

Supremo e ínfimo

En seguida se introduce el concepto de cota superior de un conjunto de números reales, el cual será de gran importancia en las secciones siguientes.

6.1 DEFINICION. Sea S un subconjunto de $\boldsymbol{R}$.

(a) Un elemento $u \in \boldsymbol{R}$ se dice que es una **cota superior** de S si $s \leq u$ para toda $s \in S$.

(b) Un elemento $w \in \boldsymbol{R}$ se dice que es una **cota inferior** de S si $w \leq s$ para toda $s \in S$.

Observe que un subconjunto $S \subseteq \boldsymbol{R}$ puede no tener una cota superior (por ejemplo, tome $S = \boldsymbol{R}$). Sin embargo, si tiene una cota superior, tiene una infinidad de ellas (ya que si u es una cota superior de S, entonces $u+n$ también es una cota superior de S para cualquier $n \in \boldsymbol{N}$). De nuevo el conjunto $S_1 = \{x \in \boldsymbol{R} : 0 < x < 1\}$ tiene a 1 como cota superior; de hecho, cualquier número $u \geq 1$ es una cota superior de S_1. Análogamente, el conjunto $S_2 = \{x \in \boldsymbol{R} : 0 \leq x \leq 1\}$ tiene las mismas cotas superiores que S_1. No obstante, observe que S_2 contiene la cota superior 1, mientras que S_1 no contiene ninguna de sus cotas superiores. (¿por qué ningún número $c < 1$ puede ser cota superior de S_1?)

Para proba que un número $u \in \boldsymbol{R}$ *no* es una cota superior de $S \subseteq \boldsymbol{R}$ se debe generar un elemento $s_0 \in S$ tal que $u < s_0$. Si $S = \emptyset$, en conjunto vacío, esto no se puede hacer. De modo que el conjunto vacío tiene la extraordinaria propiedad de que *todo* número real es una cota superior; además, todo número real es una cota inferior de $\emptyset$. Esto puede parecer ficticio; sin embargo, es una consecuencia lógica de las definiciones dadas y por lo tanto se debe aceptar.

Por cuestiones de terminología, cuando un conjunto tiene una cota superior se dirá que está **acotado por arriba** y cuando un conjunto tiene una cota inferior se dirá que está **acotado por abajo.** Si un conjunto tiene una cota superior, así como una inferior, se dice que está acotado. Si un conjunto carece de

una cota superior o de una cota inferior se dice que es **no acotado.** De modo que los cunjuntos S_1 y S_2 que antes se mencionan son acotados. Sin embargo, el subconjunto $P = \{x \in \mathbf{R} : x > 0\}$ de $\mathbf{R}$ es no acotado ya que no tiene una cota superior. Análogamente, el conjunto $\mathbf{R}$ es no acotado ya que no tiene ni cota superior ni cota inferior.

6.2 DEFINICION. Sea S un subconjunto de $\mathbf{R}$.

(a) Si S está acotado por arriba, entonces se dice que una cota superior de S es un **supremo** (o una **mínima cota superior**) de S si es menor a cualquier otra cota superior de S.

(b) Si S está acotado por abajo, entonces se dice que una cota inferior de S es un **ínfimo** (o una **máxima cota inferior**) de S si es mayor a cualquier otra cota inferior de S. (figura 6.1.)

Se explicará de otra manera: Un número $u \in \mathbf{R}$ es un supremo de un subconjunto S de $\mathbf{R}$ si satisface las siguientes condiciones:

(i) $s \leq u$ para toda $s \in S$;

(ii) si $s \leq v$ para toda $s \in S$, entonces $u \leq v$. De hecho, la condición *(i)* establece que u es una cota superior de S y *(ii)* muestra que u es menor que cualquier otra cota superior de S.

Es claro que sólo puede haber un supremo para un subconjunto dado S de $\mathbf{R}$, ya que si u_1 y u_2 son supremos de S, entonces ambos son cotas superiores de S. Como u_1 es un supremo de S y u_2 es cota superior de S, se debe tener $u_1 \leq u_2$. Siguiendo un argumento análogo se prueba que se debe tener $u_2 \leq u_1$. Por lo tanto, $u_1 = u_2$. De manera análoga se demuestra que sólo puede haber un ínfimo para un subconjunto dado S de $\mathbf{R}$. Cuando estos números existan se denotarán por medio de

$$\sup S \qquad \text{e} \qquad \inf S.$$

Con frecuencia es conveniente tener otra representación del supremo de un subconjunto $\mathbf{R}$.

6.3 LEMA. *Un número $u \in \mathbf{R}$ es el supremo de un subconjunto no vacío $S \subseteq \mathbf{R}$ si y sólo si tiene las siguientes propiedades:*

(i) No existen elementos $s \in S$ tales que $u < s$.

(ii) Si $v < u$, entonces existe un elemento $s_v \in S$ tal que $v < s_v$.

DEMOSTRACION. Suponga que u satisface *(i)* y *(ii)*. La condición *(i)* implica que u es una cota superior de S. Si v es cualquier número con

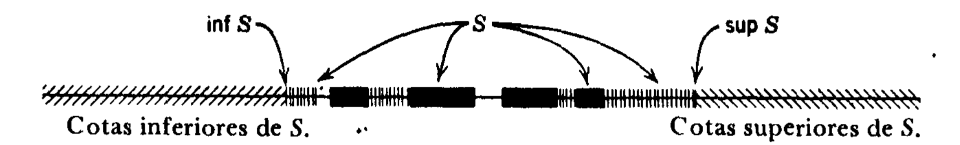


Figura 6.1 Supremos e ínfimos

$v < u$, entonces, la propiedad *(ii)* prueba que v no puede ser cota superior de S. Por lo tanto, u es el supremo de S.

A la inversa, sea u el supremo de S. Dado que u es una cota superior de S, la condición *(i)* se satisface. Si $v < u$, entonces v no es cota superior de S. Por lo tanto, existe un elemento $s_v \in S$ tal que $v < s_v$. Q.E.D.

El lector deberá convencerse por sí mismo de que el número 1 es el supremo de los dos conjuntos, S_1 y S_2 que se describen después de la definición 6.1. Observe que S_2 contiene a su supremo mientras que S_1 no contiene al suyo. De manera que *cuando se dice que un conjunto tiene un supremo no se está haciendo ninguna afirmación acerca de la contención del supremo en el conjunto.*

La siguiente afirmación es una propiedad esencial del sistema de números reales: Todo subconjunto no vacío de $\boldsymbol{R}$ que esté acotado por arriba tiene un supremo. Se hará uso frecuente de esta propiedad fundamental la cual se considera el último supuesto acerca de $\boldsymbol{R}$.

6.4 PROPIEDAD DEL SUPREMO. *Todo conjunto no vacío de números reales que tenga una cota superior tiene un supremo.*

La propiedad análoga para el ínfimo se puede deducir directamente de la propiedad del supremo.

6.5 PROPIEDAD DEL INFIMO. *Todo conjunto no vacío de números reales que tenga una cota inferior tiene un ínfimo.*

DEMOSTRACION. Sea S un conjunto acotado por abajo y sea $S_1 = \{-s : s \in S\}$ de tal manera que S_1 está acotado por arriba. La propiedad del supremo asegura que S_1 tiene un supremo u. La demostración de que $-u$ es el ínfimo de S se deja al lector. Q.E.D.

La propiedad Arquimediana†

Un resultado importante de la propiedad del supremo es que el subconjunto N de números naturales no está acotado por arriba en $\boldsymbol{R}$. Específicamente, esto significa que, dado cualquier número real x, existe un número natural n_x que es mayor que x (de otra manera, x sería una cota superior para N). En seguida se demuestra esta afirmación.

6.6 PROPIEDAD ARQUIMEDIANA. *Si $x \in \boldsymbol{R}$, existe un número natural $n_x \in \mathbf{N}$ tal que $x < n_x$.*

† Esta propiedad de $\boldsymbol{R}$ recibe este nombre por Arquímedes (287-212 A.C.), a quien se ha llamado "el más grande intelecto de la antigüedad" y que fuera uno de los fundadores del método científico.

DEMOSTRACION. Si la conclusión es falsa, entonces x es una cota superior para N. Por lo tanto, de acuerdo con la propiedad del supremo, N tiene un supremo u. Dado que x es cota superior para N, se deduce que $u \leq x$. Puesto que $u-1<u$, del lema 6.3*(ii)* se deduce que existe $n_1 \in \boldsymbol{N}$ que $u-1<n_1$. Por lo tanto, $u<n_1+1$, pero como $n_1+1 \in \boldsymbol{N}$ esto contradice el supuesto de que u sea cota superior de N. Q.E.D.

6.7 COROLARIO. *Sean y y z números reales estrictamente positivos.*
(a) Existe un número natural n tal que $ny>z$.
(b) Existe un número natural n tal que $0<1/n<z$.
(c) Existe un número natural n tal que $n-1 \leq y<n$.

DEMOSTRACION. *(a)* Dado que y y z son estrictamente positivos, $x=z/y$ también es estrictamente positivo. Sea $n \in \boldsymbol{N}$ tal que $z/y=x<n$. Entonces $z<ny$, como se quería demostrar.

(b) Sea $n \in \boldsymbol{N}$ tal que $0<1/z<n$. Entonces $0<1/n<z$.

(c) La propiedad arquimediana asegura que existen números naturales m tales que $y<m$. Sean n el mínimo de estos números naturales (vea la sección 3). Entonces $n-1 \leq y<n$. Q.E.D.

Después del teorema 5.7 se hizo ver que no existe ningún número real mínimo estrictamente positivo. El corolario 6.7*(b)* muestra que dado cualquier $z>0$ hay un número racional de la forma $1/n$ con $0<1/n<z$. Algunas veces se dice que "existen números racionales arbitrariamente pequeños de la forma $1/n$."

La existencia de $\sqrt{2}$

Como se dijo antes, una característica importante de la propiedad del supremo es que asegura la existencia de ciertos números reales. Con mucha frecuencia se hará uso de ella en estos términos. Por el momento se demostrará que esta propiedad garantiza la existencia de un número real positivo x tal que $x^2=2$; es decir, una *raíz cuadrada positiva* de 2. Este resultado viene a ser un complemento del teorema 4.7.

6.8 TEOREMA. *Existe un número positivo $x \in \boldsymbol{R}$ tal que $x^2=2$.*

DEMOSTRACION. Sea $S=\{y \in \boldsymbol{R}: 0 \leq y, y^2 \leq 2\}$. El conjunto S está acotado por arriba en 2, ya que de no ser así existiría un elemento $s \in S$ tal que $2<s$ por lo que se concluiría que $4<s^2 \leq 2$, lo que es una contradicción. Por la propiedad del supremo el conjunto S tiene un supremo, sea $x=\sup S$. Claramente se ve que $x>0$.

Se asegura que $x^2=2$. De no ser así, entonces $x^2<2$ o bien $x^2>2$. Si $x^2<2$, escójase algún $n \in \boldsymbol{N}$ de tal manera que $1/n<(2-x^2)/(2x+1)$. En este caso

$$\left(x+\frac{1}{n}\right)^2 = x^2+\frac{2x}{n}+\frac{1}{n^2} \leq x^2+\frac{2x+1}{n} < x^2+(2-x^2)=2,$$

lo que significa que $x+1/n \in S$, contrario al hecho de que x es una cota superior de S.

Si $x^2 > 2$, se elige $m \in \mathbf{N}$ tal que $1/m < (x^2-2)2x$. Dado que $x = \sup S$, existe algún $s_0 \in S$ con $x-1/m < s_0$. Pero esto implica que

$$2 < x^2 - \frac{2x}{m} < x^2 - \frac{2x}{m} + \frac{1}{m^2} = \left(x-\frac{1}{m}\right)^2 < s_0^2.$$

Por lo tanto $s_0^2 > 2$, contrario al hecho de que $s_0 \in S$.

En vista de que se han excluido las posibilidades de que $x^2 < 2$ y $x^2 > 2$, se debe tener $x^2 = 2$. Q.E.D.

Modificando ligeramente la demostración del teorema 6.8, el lector puede demostrar que si $a \geq 0$, entonces existe un número único $b \geq 0$ tal que $b^2 = a$. A b se le llama la **raíz cuadrada positiva** de a y se designa con

$$b = \sqrt{a} \qquad \text{ó} \qquad b = a^{1/2}.$$

Ahora ya se sabe que existe cuando menos un elemento irracional, a saber, $\sqrt{2}$ (la raíz cuadrada positiva de 2). De hecho, hay "más" números irracionales que números racionales en el sentido de que el conjunto de números racionales es contable mientras que el conjunto de números irracionales no es contable. (Esto se vio en la sección 3.) En seguida se demostrará que existen números irracionales arbitrariamente pequeños; este resultado complementa al corolario 6.7.

6.9 COROLARIO. Sea $\xi > 0$ *un número irracional y sea* $z > 0$. *Entonces existe un número natural* m *tal que el número irracional* ξ/m *satisface* $0 < \xi/m < z$.

DEMOSTRACION. Dado que $\xi > 0$ y $z > 0$, de los teoremas 5.6(*d*) y 5.6(*c*) se deduce que $\xi/z > 0$. Por la propiedad arquimediana existe un número natural m tal que $0 < \xi/z < m$. Por lo tanto, $0 < \xi/m < z$ y queda como ejercicio demostrar que ξ/m es irracional Q.E.D.

Ahora se demostrará que entre cualesquiera dos números reales distintos hay un número racional y un número irracional. (De hecho, hay una infinidad de ambos tipos.)

6.10 TEOREMA. *Sean* x *y* y *números reales con* $x < y$.

(*a*) *Entonces existe un número racional* r *tal que* $x < r < y$.

(*b*) *Si* $\xi > 0$ *es cualquier número irracional, entonces hay un número racional* s *tal que el número irracional* $s\xi$ *satisface* $x < s\xi < y$.

DEMOSTRACION. Sin que esto sea una simplificación se puede suponer que $0 < x$. (¿por qué?)

(a) Puesto que $y - x > 0$, del corolario 6.7*(b)* se deduce que hay un número natural m tal que $0 < 1/m < y - x$. A partir del corolario 6.7*(a)* existe un número natural k tal que

$$\frac{k}{m} = k\frac{1}{m} > x,$$

sea n el mínimo número natural que cumple la condición. Por lo tanto,

$$\frac{n-1}{m} \leq x < \frac{n}{m}.$$

Además se debe tener $n/m < y$, pues de otra manera

$$\frac{n-1}{m} \leq x < y \leq \frac{n}{m}$$

lo que implica que $y - x \leq 1/m$ contrario a la elección de m. Por lo tanto, $x < n/m < y$.

(b) Suponiendo que $0 < x < y$ y $\xi > 0$, se tiene $x/\xi < y/\xi$. Por el inciso *(a)* existe un número racional s tal que $x/\xi < s < y/\xi$. Por lo tanto, $x < s\xi < y$. (Demostrar que $s\xi$ es irracional).

Ejercicios

6.A. Demostrar que un conjunto finito no vacío de números reales tiene un supremo y un ínfimo.

6.B. Si un subconjunto S de $\boldsymbol{R}$ contiene una cota superior, entonces esta cota superior es el supremo de S.

6.C. Dar un ejemplo de un conjunto de números racionales que esté acotado pero que no tenga un supremo racional.

6.D. Dar un ejemplo de un conjunto de números irracionales que tenga un supremo racional.

6.E. Demostrar que la unión de dos conjuntos acotados es acotada.

6.F. Dar un ejemplo de una colección contable de conjuntos acotados cuya unión sea acotada y un ejemplo en donde la unión sea no acotada.

6.G. Si S es un conjunto acotado en $\boldsymbol{R}$ y si S_0 es un subconjunto no vacío de S, demostrar que

$$\inf S \leq \inf S_0 \leq \sup S_0 \leq \sup S.$$

En ocasiones es más conveniente expresar esto de otra manera. Sea $D \neq \emptyset$ y sea $f: D \to \boldsymbol{R}$ con rango acotado. Si D_0 es un subconjunto no vacío de D, entonces

$$\inf\{f(x): x \in D\} \leq \inf\{f(x): x \in D_0\} \leq \sup\{f(x): x \in D_0\} \leq \sup\{f(x): x \in D\}.$$

6.H. Sean X y Y conjuntos no vacíos y sea $f: X \times Y \to \boldsymbol{R}$ con rango acotado en $\boldsymbol{R}$. Sean

$$f_1(x) = \sup\{f(x, y): y \in Y\}, \qquad f_2(y) = \sup\{f(x, y): x \in X\}.$$

Establecer el **principio de los supremos iterados:**

$$\sup\{f(x, y) : x \in X, y \in Y\} = \sup\{f_1(x) : x \in X\}$$
$$= \sup\{f_2(y) : y \in Y\}.$$

Algunas veces se expresa esto en símbolos por medio de

$$\sup_{x,y} f(x, y) = \sup_x \sup_y f(x, y) = \sup_y \sup_x f(x, y).$$

6.I. Sean f y f_1 como en el ejercicio anterior y sea

$$g_2(y) = \inf\{f(x, y) : x \in X\}$$

Demostrar que

$$\sup\{g_2(y) : y \in Y\} \leq \inf\{f_1(x) : x \in X\}.$$

Demostrar que para desigualdad estricta puede suceder. Algunas veces esta desigualdad se expresa de la siguiente manera:

$$\sup_y \inf_x f(x, y) \leq \inf_x \sup_y f(x, y).$$

6.J. Sea X un conjunto no vacío y sea $f : X \to \boldsymbol{R}$ con rango acotado en $\boldsymbol{R}$. Si $a \in \boldsymbol{R}$, demostrar que

$$\sup\{a + f(x) : x \in X\} = a + \sup\{f(x) : x \in X\},$$
$$\inf\{a + f(x) : x \in X\} = a + \inf\{f(x) : x \in X\}.$$

6.K. Sea X un conjunto no vacío y sean f y g funciones definidas en X con rangos acotados en $\boldsymbol{R}$. Demostrar que

$$\inf\{f(x) : x \in X\} + \inf\{g(x) : x \in X\} \leq \inf\{f(x) + g(x) : x \in X\}$$
$$\leq \inf\{f(x) : x \in X\} + \sup\{g(x) : x \in X\}$$
$$\leq \sup\{f(x) + g(x) : x \in X\}$$
$$\leq \sup\{f(x) : x \in X\} + \sup\{g(x) : x \in X\}.$$

Dar ejemplos para probar que cada desigualdad puede ser absoluta.

6.L. Si $z > 0$ demostrar que existe $n \in \boldsymbol{N}$ tal que $1/2^n < z$.

6.M. Modificar la demostración del teorema 6.8 para probar que si $a > 0$, entonces el número

$$b = \sup\{y \in \boldsymbol{R} : 0 \leq y, \quad y^2 \leq a\}$$

existe y tiene la propiedad de que $b^2 = a$. Este número se denotará por medio de $\sqrt{a}$ o $a^{1/2}$, y se llama la **raíz cuadrada positiva** de a.

6.N. Emplear el ejercicio 5.P para demostrar que si $0 < a < 1$, entonces $0 < a < \sqrt{a} < 1$, mientras que si $1 < a$, entonces $1 < \sqrt{a} < a$.

Proyectos †

6.α. Si a y b son números reales estrictamente positivos y si $n \in \mathbf{N}$, ya se han definido a^n y b^n. por inducción matemática se deduce que si m, $n \in \mathbf{N}$, entonces

(i) $a^m a^n = a^{m+n}$;

(ii) $(a^m)^n = a^{mn}$;

(iii) $(ab)^n = a^n b^n$;

(iv) $a < b$ si y sólo si $a^n < b^n$.

Se adoptará el convencionalismo de que $a^0 = 1$ y $a^{-n} = 1/a^n$. En estos términos se ha definido a^x para x en $\mathbf{Z}$ y fácilmente se puede verificar que las propiedades *(i)-(ii)* siguen siendo válidas.

Se quiere definir a^x para números racionales x de tal manera que se cumplan *(i)-(ii)*. Los siguientes pasos se pueden usar como un esquema. En todos ellos se supondrá que a y b son números reales mayores de 1.

(a) Si r es un número racional dado por $r = m/n$, en donde m y n son enteros y $n > 0$, se define $S_r(a) = \{x \in \mathbf{R} : 0 \leq x^n \leq a^m\}$. Demostrar que $S_r(a)$ es un subconjunto acotado no vacío de $\mathbf{R}$ y defínase $a^r = \sup S_r(a)$.

(b) Demostrar que $z = a^r$ es la única raíz positiva de la ecuación $z^n = a^m$. *(Sugerencia:* hay una constante K tal que si $0 < \varepsilon < 1$, entonces $(1+\varepsilon)^n < 1 + K\varepsilon$. De modo que si $x^n < a^m < y^n$, existe una $\varepsilon > 0$ tal que

$$x^n(1+\varepsilon)^n < a^m < y^n/(1+\varepsilon)^n.)$$

(c) Demostrar que el valor de a^r que se dio en la parte *(a)* no depende de la representación de r de la forma m/n. Demostrar también que si r es un entero, entonces la nueva definición de a^r da el mismo valor que la anterior.

(d) Demostrar que si r, $s \in \mathbf{Q}$, entonces $a^r a^s = a^{r+s}$ y $(a^r)^s = a^{rs}$.

(e) Demostrar que $a^r b^r = (ab)^r$.

(f) Si $r \in \mathbf{Q}$, $r > 0$, entonces $a < b$ si y sólo si $a^r < b^r$.

(g) Si r, $s \in \mathbf{Q}$, entonces $r < s$ si y sólo si $a^r < a^s$.

(h) Si c es un número real tal que $0 < c < 1$, se define $c^r = (1/c)^{-r}$. Demostrar que los puntos *(d)* y *(e)* son válidos y que también es válido un resultado similar al de *(g)*, pero con la desigualdad invertida.

6.β. Una vez que a^x se ha definido para números racionales x, se desea definirla para x reales. Al hacer esto se emplearán libremente los resultados del proyecto anterior. Igual que entonces, sean a y b números reales mayores que 1. Si $u \in \mathbf{R}$, sea

$$T_u(a) = \{a^r : r \in \mathbf{Q},\ r \leq u\}.$$

Demostrar que $T_u(a)$ es un subconjunto acotado y no vacío de $\mathbf{R}$ y definir

† Se pretende que los proyectos sean un poco más alentadores para el lector; sin embargo, difieren considerablemente en cuanto a la dificultad. Se han puesto aquí estos tres proyectos (algo difíciles) porque es el lugar que lógicamente les corresponde. El lector deberá regresar a ellos más adelante, una vez que haya adquirido más experiencia con los supremos.

$$a^u = \sup T_u(a).$$

Demostrar que esta definición da el mismo resultado que la anterior cuando u es racional. Establecer las propiedades que corresponden a las afirmaciones que se dan en los incisos (d)–(g) del proyecto anterior. La función tan importante que se ha definido en $\boldsymbol{R}$ en este proyecto se llama **función exponencial** (base a). En las siguientes secciones se darán otras definiciones alternativas. Algunas veces es conveniente denotar esta función mediante el símbolo

$$\exp_a$$

y denotar su valor en el número real u por medio de $\exp_a(u)$ en vez de a^u.

6.γ. Haciendo uso de las propiedades de la función exponencial que se dan en el proyecto anterior, demostrar que $\exp_a$ es una función inyectiva con dominio $\boldsymbol{R}$ y rango $\{y \in \boldsymbol{R} : y > 0\}$. Con el supuesto fijo de que $a > 1$, esta función exponencial es estrictamente creciente en el sentido de que si $x < u$, entonces $\exp_a(x) < \exp_a(u)$. Por lo tanto, existe la función inversa con dominio $\{v \in \boldsymbol{R} : v > 0\}$ y rango $\boldsymbol{R}$. A esta función inversa se la llama **logaritmo** (base a) y se denota por

$$\log_a.$$

Demostrar que $\log_a$ es una función estrictamente creciente y que

$$\exp_a(\log_a(v)) = v \quad \textbf{para } v > 0, \qquad \log_a(\exp_a(u)) = u \quad \textbf{para } u \in \boldsymbol{R}.$$

Probar también que $\log_a(1) = 0$, $\log_a(a) = 1$, y que

$$\log_a(v) < 0 \ \text{ para } v < 1, \qquad \log_a(v) > 0 \ \text{ para } v > 1.$$

Demostrar que si $v, w > 0$, entonces

$$\log_a(vw) = \log_a(v) + \log_a(w).$$

Además, si $v > 0$ y $x \in \boldsymbol{R}$, entonces

$$\log_a(v^x) = x \log_a(v).$$

Sección 7 Cortaduras, intervalos y el conjunto de Cantor

Otro método para completar los números racionales y obtener $\boldsymbol{R}$ fue ideado por Dedekind†; se basa en el concepto de "cortadura".

7.1 DEFINICION. Se dice que un par ordenado (A, B) de subconjuntos no vacíos de $\boldsymbol{R}$ forma una cortadura si $A \cap B = \emptyset$, $A \cup B = \boldsymbol{R}$, y $a < b$ para toda $a \in A$ y toda $b \in B$.

† RICHARD DEDEKIND (1831-1916) fue alumno de Gauss. Hizo aportaciones a la teoría de números pero es mejor conocido por su trabajo en los fundamentos del sistema de números reales.

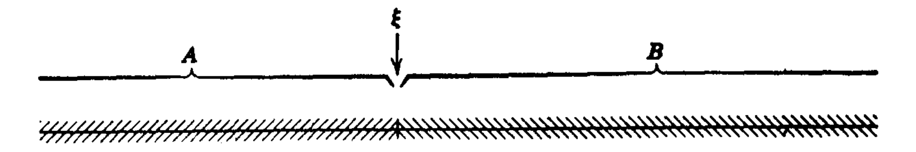


Figura 7.1 Una cortadura de Dedekind

Un ejemplo típico de una cortadura en $\boldsymbol{R}$ se obtiene para un elemento fijo $\xi \in \boldsymbol{R}$ definiendo

$$A = \{x \in \boldsymbol{R} : x \leq \xi\}, \qquad B = \{x \in \boldsymbol{R} : x > \xi\}.$$

Alternativamente se podría tomar

$$A_1 = \{x \in \boldsymbol{R} : x < \xi\}, \qquad B_1 = \{x \in \boldsymbol{R} : x \geq \xi\}.$$

Es una propiedad importante de $\boldsymbol{R}$ que toda cortadura en $\boldsymbol{R}$ esté determinada por algún número real. En seguida se establece esta propiedad.

7.2 PROPIEDAD DE CORTADURA. *Si (A, B) es una cortadura en $\boldsymbol{R}$, entonces existe un número único $\xi \in \boldsymbol{R}$ tal que $a \leq \xi$ para toda $a \in A$ y $\xi \leq b$ para toda $b \in B$.*

DEMOSTRACION. Por hipótesis, los conjuntos A y B son no vacíos. Cualquier elemento de B es cota superior de A. En consecuencia, A tiene un supremo al que se denota por ξ. Dado que ξ es una cota superior de A, entonces $a \leq \xi$ para toda $a \in A$.

Si $b \in B$, entonces, por la definición de una cortadura $a \leq b$ para toda $a \in A$. De ahí que b sea una cota superior de A y $\xi \leq b$. De esta manera se ha demostrado la existencia de un número con las propiedades establecidas.

Para ratificar la unicidad de ξ, tome $\eta \in \boldsymbol{R}$ tal que $a \leq \eta$ para toda $a \in A$ y $\eta \leq b$ para toda $b \in B$. Se deduce que η es una cota superior de A, de ahí que $\xi \leq \eta$. Si $\xi < \eta$, entonces existe un número $\zeta = (\xi + \eta)/2$ tal que $\xi < \zeta < \eta$. Ahora bien, $\zeta \in A$ o $\zeta \in B$. Si $\zeta \in A$, se contradice el hecho de que $a \leq \xi$ para toda $a \in A$. Si $\zeta \in B$, se contradice el hecho de que $\eta \leq b$ para toda $b \in B$. Por lo tanto, se debe tener $\xi = \eta$. Q.E.D.

De hecho, lo que Dedekind hizo fue esencialmente definir un número real como una cortadura en el sistema de números racionales. Este proceder permite "construir" el sistema de números reales $\boldsymbol{R}$ a partir del conjunto $\boldsymbol{Q}$ de números racionales.

Celdas e intervalos

Si $a \in \boldsymbol{R}$, entonces los conjuntos

$$\{x \in \boldsymbol{R} : x < a\}, \qquad \{x \in \boldsymbol{R} : x > a\}$$

se denominan **radios abiertos** determinados por a. Análogamente, los conjuntos.

$$\{x \in \boldsymbol{R} : x \leq a\}, \qquad \{x \in \boldsymbol{R} : x \geq a\}$$

se denominan **radios cerrados** determinados por a. Al punto a se le llama punto terminal de estos radios. A estos conjuntos con frecuencia se les asignan las siguientes notaciones

$$(-\infty, a), \qquad (a, +\infty), \qquad (-\infty, a], \qquad [a, +\infty),$$

respectivamente; en este caso $-\infty$ y $+\infty$ son simples símbolos y *no* deben considerarse elementos de $\boldsymbol{R}$.

Si $a, b \in \boldsymbol{R}$ y $a \leq b$, entonces el conjunto

$$\{x \in \boldsymbol{R} : a < x < b\}$$

se denomina **celda abierta** determinada por a y b y con frecuencia se denota por (a, b). El conjunto

$$\{x \in \boldsymbol{R} : a \leq x \leq b\}$$

se denomina **celda cerrada** determinada por a y b y se denota por $[a, b]$. Los conjuntos

$$\{x \in \boldsymbol{R} : a \leq x < b\}, \qquad \{x \in \boldsymbol{R} : a < x \leq b\}$$

se denominan **celdas semiabiertas** (o **semicerradas**) determinadas por a y b se denotan por

$$[a, b), \qquad (a, b],$$

respectivamente. A los puntos a, b se les llama los **puntos terminales** de estas celdas.

Por un **intervalo** en $\boldsymbol{R}$ se entiende un radio, una celda o bien todo $\boldsymbol{R}$. De modo que hay diez tipos diferentes de intervalos en $\boldsymbol{R}$, a saber,

$$\emptyset, \quad (-\infty, a), \quad (-\infty, a], \quad [a, b], \quad [a, b),$$
$$(a, b], \quad (a, b), \quad [b, +\infty), \quad (b, +\infty), \quad \boldsymbol{R},$$

en donde $a, b \in \boldsymbol{R}$ y $a < b$. Cinco de estos intervalos están acotados. Dos están acotados por arriba y no por abajo y dos están acotados por abajo y no por arriba.

La **celda unitaria** (o **intervalo unitario**) es el conjunto $[0, 1] = \{x \in \boldsymbol{R} : 0 \leq x \leq 1\}$. Se usará la notación clásica $\boldsymbol{I}$ para simbolizarlo.

Se dice que una sucesión de intervalos I_n, $n \in \boldsymbol{N}$, es ***nidificada*** cuando es válida la cadena

$$I_1 \supseteq I_2 \supseteq I_3 \supseteq \cdots \supseteq I_n \supseteq I_{n+1} \supseteq \cdots$$

de inclusiones. Es importante observar que una sucesión nidificada de intervalos no necesariamente debe tener un punto en común. De hecho, se de-

mostrará en un ejercicio que si $I_n = (n, +\infty)$, $n \in \mathbf{N}$, entonces la sucesión de intervalos que se obtiene es nidificada pero no tiene ningún punto en común. Análogamente, si $J_n = (0, 1/n)$, $n \in \mathbf{N}$, entonces la sucesión es nidificada pero no tiene ningún punto en común.

Sin embargo, es una propiedad muy importante de $\mathbf{R}$ en que toda sucesión nidificada de *celdas cerradas* tenga algún punto en común. En seguida se demuestra este hecho.

7.3 PROPIEDAD DE CELDAS NIDIFICADAS. *Si $n \in \mathbf{N}$, sea I_n una celda no vacía, cerrada en $\mathbf{R}$ y suponga que esta sucesión es nidificada en el sentido de que*

$$I_1 \supseteq I_2 \supseteq \cdots \supseteq I_n \supseteq \cdots.$$

Entonces existe un elemento que pertenece a todas estas celdas.

DEMOSTRACION. Suponga que $I_n = [a_n, b_n]$, en donde $a_n \leq b_n$ para toda $n \in \mathbf{N}$. Observe que $I_n \subseteq I_1$ para toda n, entonces $a_n \leq b_1$ para toda n. Por lo tanto, el conjunto $\{a_n : n \in \mathbf{N}\}$ está acotado por arriba. Sea ξ su supremo, entonces $a_n \leq \xi$ para toda n.

Se sostiene que $\xi \leq b_n$ para toda $n \in \mathbf{N}$. De no ser así, existiría alguna $m \in \mathbf{N}$ tal que $b_m < \xi$. Dado que ξ es el supremo de $\{a_n : n \in \mathbf{N}\}$ debe existir a_p tal que $b_m < a_p$. Sea q el mayor de los números naturales m y p. Puesto que $a_1 \leq a_2 \leq \cdots \leq a_n \leq \cdots$ y $b_1 \geq b_2 \geq \cdots \geq b_n \geq \cdots$ se deduce que $b_q \leq b_m < a_p \leq a_q$. Pero esto implica que $b_q < a_q$, contrario al supuesto de que $I_q = [a_q, b_q]$ en una celda cerrada *no vacía*. Por lo tanto $\xi \leq b_n$ para toda $n \in \mathbf{N}$. Como $a_n \leq \xi \leq b_n$, se deduce que $\xi \in I_n = [a_n, b_n]$ para toda $n \in \mathbf{N}$.

Q.E.D.

Observe que, con la hipótesis de 7.3, puede haber más de un elemento en común. De hecho, si se supone que $\eta = \inf\{b_n : n \in \mathbf{N}\}$, queda como ejercicio demostrar que

$$[\xi, \eta] = \bigcap_{n \in \mathbf{N}} I_n.$$

El conjunto de Cantor

En seguida se definirá un subconjunto de la celda unitaria $\mathbf{I}$ que resulta ser de gran interés y a menudo útil en la construcción de ejemplos y contraejemplos. Este conjunto se denota por $\mathbf{F}$ y se le asigna el nombre de **conjunto de Cantor** (aunque algunas veces también se le llama **conjunto ternario de Cantor** o **discontinuo de Cantor**).

Una manera de describir $\mathbf{F}$ es considerándolo como el conjunto de números reales en $\mathbf{I}$ que tienen una expansión ternaria (= base 3) usando únicamente los dígitos 0,2. Sin embargo, se decidió definirlo en otros términos.

En un sentido más preciso, F consta de aquéllos puntos en I que restan después de haber suprimido, sucesivamente, intervalos del "tercio medio".

Explícitamente: si se suprime el tercio medio abierto de I, se obtiene el conjunto

$$F_1 = [0, \tfrac{1}{3}] \cup [\tfrac{2}{3}, 1].$$

Si se suprime el tercio medio abierto de cada uno de los dos intervalos cerrados de F, se obtiene el conjunto

$$F_2 = [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1].$$

Por tanto F_2 es la unión de $4(= 2^2)$ intervalos cerrados, todos los cuales son de la forma $[k/3^2, (k+1)/3^2]$. Se puede obtener ahora el conjunto F_3 suprimiendo el tercio medio abierto de cada uno de los conjuntos. En general, si F_n ha sido construido y consta de la unión de 2^n intervalos de la forma $[k/3^n, (k+1)/3^n]$, entonces F_{n+1} se obtiene suprimiendo el tercio medio abierto de cada uno de estos intervalos. El conjunto de Cantor es lo que queda después de haber llevado a cabo este proceso para cada n en N.

7.4 DEFINICION. El conjunto de Canto F es la intersección de los conjuntos F_n, $n \in N$, que se obtienen de la supresión sucesiva de tercios medios abiertos.

A primera vista parecería que al final todo punto se suprime siguiendo este proceso. Sin embargo, éste evidentemente no es el caso ya que los puntos 0,1/3,2/3,1 pertenecen a todos los conjuntos F_n, $n \in N$, y por lo tanto n al conjunto de Cantor F. De hecho con facilidad se puede ver que hay un número infinito de puntos en F, a pesar de que F es relativamente angosto conforme a otros criterios. No es difícil demostrar que hay una cantidad no numerable de elementos de F y que los puntos de F se pueden poner en correspondencia uno a uno con los puntos de I. De ahí que el conjunto F contenga un gran número de elementos.

En seguida se dan dos criterios para los cuales F resulta ser "angosto". Primero se ve que F no contenga ningún intervalo no vacío, ya que si x pertenece a F y (a, b) es un intervalo abierto que contiene a x, entonces (a, b) contiene algunos tercios medios

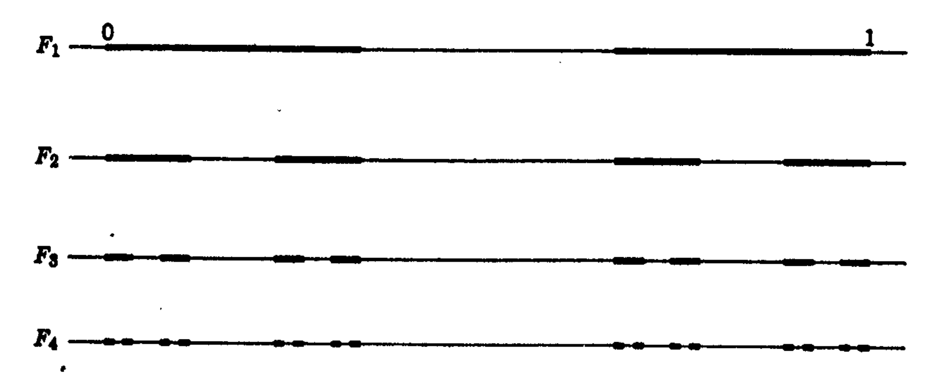


Figura 7.2 El conjunto de Cantor

que fueron suprimidos para obtener F. (¿por qué) De ahí que (a, b) no es un subconjunto del conjunto de Cantor, pero contiene una infinidad de puntos en su complemento $\mathscr{C}(\boldsymbol{F})$.

Un segundo criterio para el cual $\boldsymbol{F}$ es angosto es en relación con la "longitud". Aun cuando no es posible definir longitud para subconjuntos arbitrarios de $\boldsymbol{R}$, es fácil convencerse uno mismo de que $\boldsymbol{F}$ no puede tener longitud positiva puesto que la longitud de F_1 es $\frac{2}{3}$, la de F_2 es $\frac{4}{9}$ y, en general, la longitud de F_n es $(\frac{2}{3})^n$. Dado que $\boldsymbol{F}$ es un subconjunto de F_n, no puede tener una longitud que exceda a la de F_n. Como esto debe ser válido para cada n en $\boldsymbol{N}$, se deduce que $\boldsymbol{F}$, a pesar de ser no contable, no puede tener longitud positiva.

A pesar de lo extraño que puede parecer el conjunto Cantor, observa relativamente un buen comportamiento en muchos sentidos. Proporciona una idea de lo complicado que pueden ser algunos subconjuntos de $\boldsymbol{R}$ y lo poco que nos guía la intuición. También sirve como una prueba para los conceptos que se darán en secciones posteriores y cuyos significados no se pueden comprender del todo en términos de intervalos y de otros subconjuntos muy elementales.

Modelos para *R*

En las secciones 4-6 se ha definido $\boldsymbol{R}$ axiomáticamente en el sentido de que se ha proporcionado una lista de algunas propiedades que se supone tiene el conjunto. Esta definición hace que surja una pregunta en cuanto a si el conjunto en realidad existe y hasta qué punto está determinado de manera única. Aunque no se habrán de contestar estas preguntas, sin duda son convenientes algunas aclaraciones acerca de ellas.

La existencia de un conjunto que sea un campo ordenado completo se puede demostrar por simple construcción. Si uno se siente lo suficientemente familiarizado con el campo racional $\boldsymbol{Q}$, puede definir a los números reales como subconjuntos especiales de $\boldsymbol{Q}$ y definir suma, multiplicación y relaciones de orden entre estos subconjuntos de tal manera que se obtenga un campo ordenado completo. Existen dos procedimientos clásicos que se usan para hacer esto: uno es el método de Dedekind de "cortaduras" que se analiza en el libro de Rudin citado en la bibliografía. El otro es el método de Cantor de "sucesiones de Cauchy" que se analiza en el libro de Hamilton y Landin.

En el último párrafo se ha asegurado que es posible construir un modelo de $\boldsymbol{R}$ a partir de Q (cuando menos de dos maneras diferentes). También es posible construir un modelo de $\boldsymbol{R}$ a partir del conjunto N de números naturales y es lo que a menudo se toma como punto de partida por aquellos que, como Kronecker†, consideran a los números naturales dados por Dios. Sin embargo, puesto que hasta el conjunto de números naturales tiene algunas sutilezas (como es la propiedad del buen orden), se cree que el procedimiento más

† LEOPOLD KRONECKER (1823-1891) estudió con Dirichlet en Berlín y con Kummer en Bonn. Después de haber hecho una fortuna, antes de los treinta años, regresó a las matemáticas. Es conocido por su trabajo en álgebra y teoría de números y por su oposición personal a las ideas de Cantor acerca de la teoría de los conjuntos.

satisfactorio es seguir el proceso de construir primero el conjunto N a partir de conceptos primitivos de teoría de conjuntos, después construir el conjunto de enteros, luego el campo Q de racionales y por último el conjunto $\boldsymbol{R}$. Este procedimiento no es especialmente difícil de seguir y es edificante; sin embargo, es un poco largo. Dado que se proporciona con detalle en el libro de Hamilton y Landin, no se dará aquí.

A partir de las observaciones ya hechas, es claro que se pueden construir campos ordenados completos de distintas maneras. De modo que no se puede decir que hay un campo ordenado completo *único*. En cierto sentido, de todos los métodos de construcción sugeridos con anterioridad se derivan campos ordenados completos que son "isomorfos." (Esto significa que si $\boldsymbol{R}_1$ y $\boldsymbol{R}_2$ son campos ordenados completos obtenidos a partir de estas construcciones, entonces existe una aplicación uno a uno φ de $\boldsymbol{R}_1$ sobre $\boldsymbol{R}_2$ tal que *(i)* φ manda un elemento racional de $\boldsymbol{R}_1$ hacia el elemento racional correspondiente de $\boldsymbol{R}_2$, *(ii)* φ manda $a+b$ hacia $\varphi(a)+\varphi(b)$, *(iii)* φ manda ab hacia $\varphi(a)\varphi(b)$ y *(iv)* φ manda un elemento positivo de $\boldsymbol{R}_1$ hacia un elemento positivo de $\boldsymbol{R}_2$.) Con base en la teoría de conjuntos elemental, se puede proporcionar un argumento en el que se demuestre que cualquier par de campos ordenados completos es isomorfo en el sentido recien descrito. El que esté argumentado se pueda formalizar a partir de un sistema de lógica dado depende de las reglas de inferencia que se empleen en el sistema. De modo que la pregunta acerca de hasta qué grado el sistema de números reales se puede considerar determinado de manera única es un problema de lógica bastante delicado. Sin embargo, para nuestros propósitos esta unicidad (o falta de ella) no es importante, ya que se puede escoger algún campo ordenado completo en particular como modelo para el sistema de números reales.

Ejercicios

7.A. Si (A, B) es una cortadura en $\boldsymbol{R}$, demostrar que $\sup A = \inf B$.

7.B. Si las cortaduras (A, B) y (A', B') determinan en los números reales ξ y ξ', respectivamente, demostrar que $\xi < \xi'$ implica que $A \subseteq A'$, $A \neq A'$.

7.C. ¿Es verdad el inverso del ejercicio anterior?

7.D. Sea $A = \{x \in \boldsymbol{R} : x \leq 0 \text{ y } x^2 \leq 2\}$ y $B = \{x \in \boldsymbol{R} : x > 0 \text{ y } x^2 > 2\}$. Demostrar que (A, B) es una cortadura en $\boldsymbol{R}$.

7.E. Sea $I_n = (n, +\infty)$ para $n \in \boldsymbol{N}$. Demostrar que la sucesión de intervalos es nidificada pero que no hay ningún punto en común.

7.F. Sea $J_n = (0, 1/n)$ para $n \in \boldsymbol{N}$. Demostrar que esta sucesión de intervalos es nidificada pero que no hay ningún punto en común.

7.G. Si $I_n = [a_n, b_n]$, $n \in \boldsymbol{N}$, es una sucesión nidificada de celdas cerradas, demostrar que

$$a_1 \leq a_2 \leq \cdots \leq a_n \leq \cdots \leq b_m \leq \cdots \leq b_1.$$

Si se toma $\xi = \sup\{a_n : n \in \boldsymbol{N}\}$ y $\eta = \inf\{b_m : m \in \boldsymbol{N}\}$, demostrar que $[\xi, \eta] = \bigcap_{n \in \boldsymbol{N}} I_n$.

7.H. Demostrar que todo número en el conjunto de Cantor tiene una expansión ternaria (= base 3) usando únicamente los dígitos 0,2.

7.I. Demostrar que la colección de puntos terminales "por la derecha" en $\boldsymbol{F}$ es numerable. Demostrar que si todos estos puntos terminales se suprimen de $\boldsymbol{F}$, entonces lo que queda se puede poner en correspondencia uno a uno con todo $[0, 1)$. Deducir que el conjunto $\boldsymbol{F}$ no es contable.

7.J. Todo intervalo abierto (a, b) que contiene un punto de $\boldsymbol{F}$ también contiene a todo un conjunto "tercio medio" que pertenece a $\mathscr{C}(\boldsymbol{F})$. Por lo tanto, $\boldsymbol{F}$ no contiene ningún intervalo abierto no vacío.

7.K. Al suprimir conjuntos cuya longitud es decreciente, demostrar que se puede construir un conjunto "a la manera de Cantor" que tenga longitud positiva. ¿Qué tan grande se puede hacer la longitud de este conjunto?

7.L. Demostrar que $\boldsymbol{F}$ no es la unión de un conjunto contable de intervalos cerrados.

II LA TOPOLOGIA DE ESPACIOS CARTESIANOS

Las secciones del capítulo I se dedicaron al desarrollo de las propiedades algebraicas, las propiedades de orden y la propiedad de completación del sistema de números reales. Dichas propiedades se emplearán en forma constante en este capítulo y en otros posteriores.

Aunque sería posible pasar de inmediato al análisis de sucesiones de números reales y funciones continuas reales, es preferible atrasar un poco el estudio de estos temas. De hecho, se intercalarán aquí las definiciones de espacio vectorial, espacio normado y espacio de producto interno, ya que estos conceptos se entienden con facilidad y aparecen en todo el análisis (sin mencionar sus aplicaciones a la geometría, la física, la ingeniería, la economía, etc.). Desde luego, los espacios cartesianos $\boldsymbol{R}^p$ serán de especial interés. Por fortuna, la idea que se tiene de $\boldsymbol{R}^2$ y $\boldsymbol{R}^3$ generalmente persiste, sin mucho cambio, en el espacio $\boldsymbol{R}^p$. El conocimiento de estos espacios es útil para analizar espacios más generales.

Sección 8 Espacios vectoriales y cartesianos

Un "espacio vectorial" es un conjunto en el que se pueden sumar dos elementos y se puede multiplicar un elemento por un número real de tal manera que ciertas propiedades conocidas sean válidas. Ahora, con más precisión, se tiene lo siguiente:

8.1 DEFINICION. Un **espacio vectorial** es un conjunto V (a cuyos elementos se les da el nombre de **vectores**) provisto de dos operaciones binarias llamadas **suma vectorial** y **multiplicación escalar.**

Si $x, y \in V$ hay en V un elemento $x + y$ llamado la **suma vectorial** de x y y. Esta operación de suma vectorial satisface las siguientes propiedades:

(A1) $x + y = y + x$ para todas x, y en V;

(A2) $(x + y) + z = x + (y + z)$ para todas x, y, z en V;

(A3) existe un elemento O en V tal que $0+x=x$ y $x+0=x$ para toda x en V;

(A4) dada x en V hay un elemento $-x$ en V tal que $x+(-x)=0$ y $(-x)+x=0$. Si $a \in \boldsymbol{R}$ y $x \in V$ hay un elemento ax en V llamado el **múltiplo** de a y de x. Esta operación de multiplicación escalar satisface las siguientes propiedades:

(M1) $1x=x$ para toda $x \in V$;

(M2) $a(bx)=(ab)x$ para todas $a, b \in \boldsymbol{R}$ y $x \in V$;

(D) $a(x+y)=ax+ay$ y $(a+b)x=ax+by$ para todos los reales $a, b \in \boldsymbol{R}$ y $x, y \in V$.

En seguida se darán algunos ejemplos, elementales pero importantes, de espacios vectoriales.

8.2 EJEMPLOS. *(a)* El sistema de números reales es un espacio vectorial en que las operaciones de suma y de multiplicación escalar son la suma y la multiplicación ordinarias de los números reales.

(b) Se denotará al producto cartesiano $\boldsymbol{R} \times \boldsymbol{R}$ por medio de $\boldsymbol{R}^2$. Entonces, $\boldsymbol{R}^2$ consta de todos los pares ordenados (x_1, x_2) de números reales. Si se definen suma vectorial y multiplicación escalar por

$$(x_1, x_2)+(y_1, y_2)=(x_1+y_1, x_2+y_2),$$
$$a(x_1, x_2)=(ax_1, ax_2),$$

entonces fácilmente se puede comprobar que las propiedades de la definición 8.1 se satisfacen. En este caso $[0=(0,0)$ y $-(x_1, x_2)=(-x_1, -x_2).]$ Por lo tanto, $\boldsymbol{R}^2$ es un espacio vectorial bajo estas operaciones.

(c) Sea $p \in \boldsymbol{N}$ y sea $\boldsymbol{R}^p$ la colección de todas las "p-adas" ordenadas

$$(x_1, x_2, \ldots, x_p)$$

con $x_i \in \boldsymbol{R}$ para $i=1, \ldots, p$. Si se definen suma vectorial y multiplicación escalar por

$$(x_1, x_2, \ldots, x_p)+(y_1, y_2, \ldots, y_p)=(x_1+y_1, x_2+y_2, \ldots, x_p+y_p)$$
$$a(x_1, x_2, \ldots, x_p)=(ax_1, ax_2, \ldots, ax_p),$$

entonces con facilidad se puede comprobar que $\boldsymbol{R}^p$ es un espacio vectorial bajo estas operaciones. Se puede ver que $[0=(0,0,\ldots,0)$ y $-(x_1, x_2, \ldots, x_p)=(-x_1, -x_2, \ldots, -x_p).]$.

(d) Sea S cualquier conjunto y denote $\boldsymbol{R}^S$ a la colección de todas las funciones u con dominio S y rango en $\boldsymbol{R}$. (De donde, $\boldsymbol{R}^S$ es la colección de todas las funciones de valor real definidas en S.) Si se define $u+v$ y au por

$$(u+v)(s)=u(s)+u(s),$$
$$(au)(s)=au(s),$$

para toda $s \in S$, se puede comprobar con facilidad que $\mathbf{R}^S$ es un espacio vectorial bajo estas operaciones. Es este caso O es la función idénticamente igual a cero y $-\boldsymbol{u}$ es la función cuyo valor en $s \in S$ es $-\boldsymbol{u}(s)$.].

En secciones posteriores se encontrarán muchos otros espacios vectoriales.

Por lo general, se escribirá $x - y$ en vez de $x + (-y)$.

Productos internos y normas

El lector observará que la multiplicación escalar en un espacio vectorial V es una función con dominio $\boldsymbol{R} \times V$ y rango V. Muchos espacios vectoriales también están provistos de una función, con dominio $V \times V$ y rango $\boldsymbol{R}$, que resulta ser importante

8.3 DEFINICION. Si V es un espacio vectorial, entonces un **producto interno** (o **producto puntual**) es una función de $V \times V$ a $\boldsymbol{R}$, denotada por $(x, y) \mapsto x \cdot y$, que satisface las siguientes propiedades:

(i) $x \cdot x \geq 0$ para toda $x \in V$;

(ii) $x \cdot x = 0$ si y sólo si $x = 0$;

(iii) $x \cdot y = y \cdot x$ para todas $x, y \in V$;

(iv) $x \cdot (y + z) = x \cdot y + x \cdot z$ y $(x + y) \cdot z = x \cdot z + y \cdot z$ para todas $x, y, z \in V$;

(v) $(ax) \cdot y = a(x \cdot y) = x \cdot (ay)$ para todas $a \in \boldsymbol{R}$, y $x, y \in V$.

Un espacio vectorial en el que se ha definido un producto interno se llama **espacio de producto interno**.

Es posible que productos internos distintos se definan en un mismo espacio vectorial (cf. ejercicio 8.D).

8.4 EJEMPLOS. *(a)* La multiplicación ordinaria en $\boldsymbol{R}$ satisface las propiedades anteriores; por tanto, $\boldsymbol{R}$ es un espacio de producto interno.

(b) En $\boldsymbol{R}^2$, se define

$$(x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 + x_2 y_2.$$

Resulta fácil comprobar que esto define prcisamente a un producto interno en $\boldsymbol{R}^2$.

(c) En $\boldsymbol{R}^p$, se define

$$(x_1, x_2, \ldots, x_p) \cdot (y_1, y_2, \ldots, y_p) = x_1 y_1 + x_2 y_2 + \cdots + x_p y_p.$$

Resulta fácil comprobar que esto define a un producto interno en $\boldsymbol{R}^p$.

8.5 DEFINICION. Si V es un espacio vectorial, entonces una **norma en** V es una función de V a $\boldsymbol{R}$ denotada por $x \mapsto \|x\|$ que satisface las siguientes propiedades:

(i) $\|x\| \geq 0$ para toda $x \in V$;
(ii) $\|x\| = 0$ si y sólo si $x = 0$;
(iii) $\|ax\| = |a|\,\|x\|$ para toda $a \in \mathbf{R}$, $x \in V$;
(iv) $\|x + y\| \leq \|x\| + \|y\|$ para toda x, $y \in V$.

Un espacio vectorial en el que se ha definido una norma recibe el nombre de **espacio normado.**

En los ejercicios se verá que un mismo espacio vectorial puede tener varias normas interesantes.

8.6 EJEMPLOS. *(a)* La función valor absoluto en $\mathbf{R}$ satisface las propiedades que aparecen en 8.5.

(b) En $\mathbf{R}^2$, se define

$$\|(x_1, x_2)\| = (x_1{}^2 + x_2{}^2)^{1/2}.$$

Las propiedades *(i)*, *(ii)* y *(iii)* son muy fáciles de comprobar. La propiedad *(iv)* es un poco más complicada.

(c) En $\mathbf{R}^p$, se define

$$\|(x_1, x_2, \ldots, x_p)\| = (x_1{}^2 + x_2{}^2 + \cdots + x_p{}^2)^{1/2}.$$

De nuevo las propiedades *(i)*, *(ii)* y *(iii)* son fáciles.

En seguida se dará un teorema en el que se afirma que un producto interno siempre se puede usar para definir una norma de una manera muy natural.

8.7 TEOREMA. *Sea V un producto interno y defínase* $\|x\|$ *por medio de*

$$\|x\| = \sqrt{x \cdot x} \quad para \quad x \in V.$$

Entonces $x \mapsto \|x\|$ *es una norma en V y satisface la propiedad*

$$(*) \qquad x \cdot y \leq \|x\|\,\|y\|.$$

Además, si x y y son distintos de cero, entonces la igualdad se satisface en (x) si y sólo si hay algún número real estrictamente positivo tal que $x = cy$.

DEMOSTRACION. Dado que $x \cdot x \geq 0$ para toda $x \in V$, la raíz cuadrada de $x \cdot x$ existe, de manera que $\|x\|$ está bien definido. Las tres primeras propiedades de la norma son consecuencia directa de 8.3 *(i)*, *(ii)* y *(v)*. Para demostrar *(x)*, sean $a, b \in \mathbf{R}$, $x, y \in V$, y $z = ax - by$. Empleando las propiedades 8.3 *(i)*, *(iii)*, *(iv)* y *(v)* se obtiene

$$0 \leq z \cdot z = a^2\, x \cdot x - 2ab\, x \cdot y + b^2\, y \cdot y.$$

Tome ahora $a = \|y\|$ y $b = \|x\|$, para obtener

i077 (1587x2511x2 pcx)

$$0 \leq \|y\|^2 \|x\|^2 - 2\|y\|\,\|x\|\, x \cdot y + \|x\|^2 \|y\|^2$$
$$= 2\|x\|\,\|y\|\,(\|x\|\,\|y\| - x \cdot y).$$

Por lo tanto, la desigualdad *(x)* se satisface. Si $x = cy$ con $c > 0$, entonces $\|x\| = c\,\|y\|$ por lo que

$$x \cdot y = (cy) \cdot y = c(y \cdot y) = c\,\|y\|^2$$
$$= \|x\|\,\|y\|$$

de modo que la igualdad es válida en *(x)*. De manera inversa, si $x \cdot y = \|x\|\,\|y\|$ el cálculo hecho en el párrafo anterior muestra que $z = \|y\|\,x - \|x\|\,y$ tiene la propiedad de que $z \cdot z = 0$. Por lo tanto, $z = 0$ y puesto que x y y son vectores distintos de cero, se puede tomar $c = \|x\|/\|y\|$. Para probar 8.5 *(iv)* se usa (*) y se demuestra que

$$\|x + y\|^2 = (x + y) \cdot (x + y)$$
$$= x \cdot x + x \cdot y + y \cdot x + y \cdot y$$
$$= \|x\|^2 + 2(x \cdot y) + \|y\|^2$$
$$\leq \|x\|^2 + 2\|x\|\,\|y\| + \|y\|^2$$
$$\leq (\|x\| + \|y\|)^2,$$

por lo tanto, se deduce que $\|x + y\| \leq \|x\| + \|y\|$ para todas $x, y \in V$.

La demostración del siguiente corolario se deja como ejercicio.

8.8 COROLARIO. *Si x, y son elementos de V, entonces*

(**) $$|x \cdot y| \leq \|x\|\,\|y\|.$$

*Además, si $y \neq 0$, la igualdad es válida en (*9) si y sólo si hay un número real c tal que $x = cy$.*

Las dos desigualdades (*9) y (**9) reciben el nombre de **desigualdad de Schwarz** o **desigualdad de Cauchy-Bunyakovskii-Schwarz.**† Estas desigualdades se usarán con frecuencia posteriormente. A la desigualdad 8.5 *(iv)* se le llama **desigualdad del triángulo.** Se deja al lector demostrar que

†AUGUSTIN-LOUIS CAUCHY (1789-1857) fue el fundador del análisis moderno, pero también hizo importantes aportaciones en otras áreas de las matemáticas. Prestó sus servicios como ingeniero al mando de Napoleón, siguió a Carlos X en el exilio, que él mismo se impuso, y fue expulsado de su puesto en el "Colegio de Francia" durante los años de la monarquía de Julio por no prestar un juramento de lealtad. A pesar de sus actividades políticas y religiosas encontró tiempo para escribir 789 trabajos acerca de matemáticas.

VICTOR BUNYAKOVSKII (1804-1889), profesor en San Petersburgo, estableció una generalización de la desigualdad de Cauchy para integrales en 1859-1859. Su contribución pasó inadvertida para los escritores occidentales y más tarde Schwarz la descubrió independientemente.

HERMANN AMANDUS SCHWARZ (1843-1921) fue alumno y sucesor de Weierstrass en Berlín. Hizo numerosas aportaciones, especialmente al análisis complejo.

$$| \|x\| - \|y\| | \leq \|x \pm y\| \leq \|x\| + \|y\|,$$

para cualesquiera x, y en un espacio normado.

El espacio cartesiano $\boldsymbol{R}^p$

El **espacio cartesino real p-dimensional** es el conunto $\boldsymbol{R}^p$ provisto de la suma vectorial y la multiplicación escalar definidas en el ejemplo 8.2*(c)* y del producto interno definido en el ejemplo 8.4*(c)*. Como ya se ha visto, este producto interno da origen a la norma

$$\|(x_1, x_2, \ldots, x_p)\| = \sqrt{x_1{}^2 + x_2{}^2 + \cdots + x_p{}^2}.$$

Los números reales $x_1, x_2, \ldots, x_p$ reciben el nombre de **primera, segunda, ..., p-esima coordenadas** (o **componentes**) del vector $x = (x_1, x_2, \ldots, x_p)$.

En $\boldsymbol{R}^p$, el número real $\|x\|$ se puede considerar como la "longitud" de x o como la distancia de x a 0. Generalizando, $\|x - y\|$se le considera como la distancia de x a y. Con esta interpretación, la propiedad 8.5 *(ii)* asegura que la distancia de x a y es cero si y sólo si $x = y$. La propiedad 8.5*(iii)*, con $a = -1$, asegura que $\|x - y\| = \|y - x\|$, lo que significa que la distancia de x a y es igual a la distancia de y a x. La desigualdad del triángulo implica que

$$\|x - y\| \leq \|x - z\| + \|z - y\|,$$

lo que significa que la distancia de x a y no es mayor a la suma de la distancia de x a z y la distancia de z a y.

8.9 DEFINICION. Sean $x \in \boldsymbol{R}^p$ y $r > 0$. Entonces, al conjunto $\{y \in \boldsymbol{R}^p : \|x - y\| < r\}$ se le llama la **bola abierta** con **centro** x y radio r. Al conjunto $\{y \in \boldsymbol{R}^p : \|x - y\| \leq r\}$ se le llama la **bola cerrada** con **centro** x y **radio** r. Al conjunto $\{y \in \boldsymbol{R}^p : \|x - y\| = r\}$ se le llama la **esfera** con **centro** x y **radio** r.

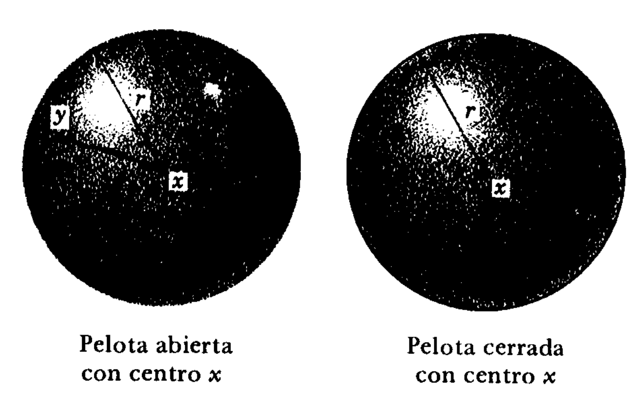
La descripción de una bola depende de la norma. En los ejercicios se verá que algunas bolas no son muy "redondas".

A menudo es conveniente que haya relaciones entre la norma de un vector en $\boldsymbol{R}^p$ y la magnitud de sus componentes.

8.10 TEOREMA. *Si $x = (x_1, x_2, \ldots, x_p)$ es cualquier elemento de $\boldsymbol{R}^p$, entonces*

$$|x_i| \leq \|x\| \leq \sqrt{p} \sup \{|x_1|, |x_2|, \ldots, |x_p|\}.$$

DEMOSTRACION. Dado que $\|x\|^2 = x_1{}^2 + x_2{}^2 + \cdots + x_p{}^2$, es claro que $|x_i| \leq \|x\|$ para toda i. Análogamente, si $M = \sup \{|x_1|, |x_2|, \ldots, |x_p|\}$, entonces $\|x\|^2 \leq pM^2$, por lo que $\|x\| \leq \sqrt{p}\, M$.



Pelota abierta con centro x

Pelota cerrada con centro x

Figura 8.1

La desigualdad que se acaba de establecer asegura, en forma cuantitativa, que si la norma de x es pequeña, entonces las longitudes de sus componentes son pequeñas. Se asegura también lo inverso.

Ejercicios

8.A. Si V es un espacio vectorial y si $x + z = x$ para algunas x y z en V, demostrar que $z = 0$. Por lo tanto, el elemento cero en V es único.

8.B. Si $x + y = 0$ para algunas x y y en V, demostrar que $y = -x$.

8.C. Sea $S = \{1, 2, \ldots, p\}$, para alguna $p \in \boldsymbol{N}$. Demostrar que el espacio vectorial $\boldsymbol{R}^S$ es "esencialmente el mismo" que el espacio $\boldsymbol{R}^p$.

8.D. Si w_1 y w_2 son estrictamente positivos, demostrar que la definición

$$(x_1, x_2) \cdot (y_1, y_2) = x_1 y_1 w_1 + x_2 y_2 w_2,$$

da un producto interno en $\boldsymbol{R}^2$. Generalizar esto para $\boldsymbol{R}^p$.

8.E. La definición

$$(x_1, x_2) \cdot (y_1, y_2) = x_1 y_1$$

no es un producto interno en $\boldsymbol{R}^2$. ¿Por qué?

8.F. Si $x = (x_1, x_2, \ldots, x_p) \in \boldsymbol{R}^p$, defínase $\|x\|_1$ por medio de

$$\|x\|_1 = |x_1| + |x_2| + \cdots + |x_p|.$$

Demostrar que $x \mapsto \|x\|_1$ es una norma en $\boldsymbol{R}^p$.

8.G. Si $x = (x_1, x_2, \ldots, x_p) \in \boldsymbol{R}^p$, defínase $\|x\|_\infty$ por medio de

$$\|x\|_\infty = \sup \{|x_1|, |x_2|, \ldots, |x_p|\}.$$

Demostrar que $x \mapsto \|x\|_\infty$ es una norma en $\boldsymbol{R}^p$.

8.H. En el conjunto $\boldsymbol{R}^2$, describir los conjuntos

$$S_1 = \{x \in \boldsymbol{R}^2 : \|x\|_1 < 1\}, \qquad S_\infty = \{x \in \boldsymbol{R}^2 : \|x\|_\infty < 1\}.$$

8.I. Si x, $y \in \boldsymbol{R}^p$, la norma definida en 8-4(*c*) satisface la **igualdad del paralelogramo:**

$$\|x+y\|^2+\|x-y\|^2=2(\|x\|^2+\|y\|^2).$$

Demostrar esto y probar que se puede interpretar diciendo que la suma de los cuadrados de las longitudes de los cuatro lados de un paralelogramo es igual a la suma de los cuadrados de las diagonales

8.J. Demostrar que las normas definidas en los ejercicios 8.F y 8.G no satisfacen la igualdad del paralelogramo.

8.K. Demostrar que existen constantes positivas a, b tales que

$$a\,\|x\|_1 \le \|x\| \le b\,\|x\|_1 \quad \text{para toda } x \in \boldsymbol{R}^p.$$

Encontrar la mayor constante a y la menor constante b con esta propiedad.

8.L. Demostrar que existen constantes positivas a, b tales que

$$a\,\|x\|_1 \le \|x\|_\infty \le b\,\|x\|_1 \quad \text{para toda } x \in \boldsymbol{R}^p.$$

Encontrar la mayor constante a y la menor constante b con esta propiedad.

8.M. Si x, y pertenecen a $\boldsymbol{R}^p$, ¿Es verdad qué

$$|x\cdot y| \le \|x\|_1\,\|y\|_1 \qquad \text{y} \qquad |x\cdot y| \le \|x\|_\infty\,\|y\|_\infty\,?$$

8.N. Si x, y pertenecen a $\boldsymbol{R}^p$, ¿Es verdad que la relación

$$\|x+y\|=\|x\|+\|y\|$$

es válida si y sólo si $x = cy$ ó $y = cx$ con $c \ge 0$?

8.O. Perteneciendo x, y a $\boldsymbol{R}^p$, ¿Es verdad que la relación

$$\|x+y\|_\infty=\|x\|_\infty+\|y\|_\infty$$

es válida si y sólo si $x = cy$ ó $y = cx$ con $c \ge 0$?

8.P. Si x, y pertenecen a $\boldsymbol{R}^p$, entonces

$$\|x+y\|^2=\|x\|^2+\|y\|^2$$

es válido si y sólo si $x \cdot y = 0$. En este caso se dice que x y y son **ortogonales** o **perpendiculares.**

8.Q. Un subconjunto K de $\boldsymbol{R}^p$ se dice que es **convexo** si, siempre que x, y pertenezcan a K y t sea un número real tal que $0 \le t \le 1$, entonces el punto

$$(1-t)x+ty=x+t(y-x)$$

también pertenece a K. Interpretar esta condición geométricamente y demostrar que los subconjuntos

$$K_1=\{x\in\boldsymbol{R}^2:\|x\|\le 1\},$$
$$K_2=\{(\xi,\eta)\in\boldsymbol{R}^2:0<\xi<\eta\},$$
$$K_3=\{(\xi,\eta)\in\boldsymbol{R}^2:0\le\eta\le\xi\le 1\},$$

son convexos pero que el subconjunto

$$K_4 = \{x \in \mathbf{R}^2 : \|x\| = 1\}$$

no lo es.

8.R. La intersección de cualquier colección de subconjuntos convexos de $\mathbf{R}^p$ es convexa. La unión de dos subconjuntos convexos de $\mathbf{R}^p$ puede no ser convexa.

8.S. Si M es cualquier conjunto, entonces a una función $d: M \times M \to \mathbf{R}$ se la llama una **métrica** en M si satisface :

(i) $d(x, y) \geq 0$ para todas en x, y en M,

(ii) $d(x, y) = 0$ si y sólo si $x = y$;

(iii) $d(x, y) = d(y, x)$ para todas x, y en M;

(iv) $d(x, y) \leq d(x, z) + d(z, y)$ para todas x, y, z en M.

Demostrar que si $x \mapsto \|x\|$ es alguna norma en un espacio vectorial V y si se define d por medio de $d(x, y) = \|x - y\|$, para $x, y \in V$, entonces d es una métrica en V.

8.T. Suponga que d es una métrica en un conjunto M. Usando la definición 8.9 como modelo, definir una bola abierta con centro $x \in M$ y radio r. Representar a los conjuntos S_1 y S_∞ del ejercicio 8.H como bolas abiertas en $\mathbf{R}^2$ con respecto a dos métricas diferentes. Interpretar el ejercicio 8.K diciendo que una bola con centro 0 respecto de la métrica d_2 (obtenida de la norma en 8.6*(b)*) contiene y está contenida en bolas con centro 0 respecto de la métrica d_1 que se obtuvo de $\| \quad \|_1$. Hacer interpretaciones análogas del ejercicio 8.L y del teorema 8.10.

8.U. Sea M cualquier conjunto y defina a d en $M \times M$ de manera que

$$d(x, y) = \begin{cases} 0 & \text{si} \quad x = y, \\ 1 & \text{si} \quad x \neq y. \end{cases}$$

Demostrar que d da una métrica en M en el sentido que se definió en el ejercicio 8.S. Si x es cualquier punto en M, entonces la bola abierta con centro x y radio 1 (con respecto a la métrica d) consta precisamente de un punto. Sin embargo, la bola abierta con centro x y radio 2 (con respecto a d) consta de todo M. A esta métrica d algunas veces se la llama la **métrica discreta** en el conjunto M.

Proyectos

8.α. En este proyecto se establecen algunas desigualdades importantes.

(a) Sean a y b números reales positivos. Demostrar que

$$ab \leq (a^2 + b^2)/2,$$

y que la igualdad es válida si y sólo si $a = b$. *(Sugerencia:* considere $(a-b)^2$.)

(b) Sean a_1 y a_2 números reales positivos. Demostrar que

$$\sqrt{a_1 a_2} \leq (a_1 + a_2)/2$$

y que la igualdad es válida si y sólo si $a_1 = a_2$.

(c) Sean $a_1, a_2, \ldots, a_m$ $m = 2^n$ números reales positivos. Demostrar que

$$(*) \qquad (a_1 a_2 \cdots a_m)^{1/m} \leq (a_1 + a_2 + \cdots + a_m)/m$$

y que la igualdad es válida si y sólo si $a_1 = \cdots = a_m$.

(d) Demostrar que la desigualdad (*) entre la media geométrica y la media aritmética es válida aun cuando m no sea potencia de 2. *(Sugerencia:* si $2^{n-1} < m < 2^n$, sea $b_j = a_j$ para $j = 1, \ldots, m$ sea

$$b_j = (a_1 + a_2 + \cdots + a_m)/m$$

para $j = m+1, \ldots, 2^n$. Aplicar ahora el inciso *(c)* a los números $b_1, b_2, \ldots, b_{2^n}$.)

(e) Sean $a_1, a_2, \ldots, a_n$ y $b_1, b_2, \ldots, b_n$ dos conjuntos de número reales. Demostrar la **igualdad de Lagrange**†

$$\left\{\sum_{j=1}^{n} a_j b_j\right\}^2 = \left\{\sum_{j=1}^{n} a_j^2\right\}\left\{\sum_{k=1}^{n} b_k^2\right\} - \frac{1}{2}\sum_{j,k=1}^{n} (a_j b_k - a_k b_j)^2.$$

(Sugerencia: intentar lo primero para $n = 2$ y $n = 3$.)

(f) Usar la parte *(e)* para probar la **desigualdad de Cauchy**

$$\left\{\sum_{j=1}^{n} a_j b_j\right\}^2 \le \left\{\sum_{j=1}^{n} a_j^2\right\}\left\{\sum_{k=1}^{n} b_k^2\right\}.$$

Probar que la igualdad es válida si y sólo si los conjuntos ordenados $(a_1, a_2, \ldots, a_n)$ y $(b_1, b_2, \ldots, b_n)$ son proporcionales.

(g) Usar la parte *(f)* para probar la **desigualdad del triángulo**

$$\left\{\sum_{j=1}^{n} (a_j + b_j)^2\right\}^{1/2} \le \left\{\sum_{j=1}^{n} a_j^2\right\}^{1/2} + \left\{\sum_{j=1}^{n} b_j^2\right\}^{1/2}$$

8.β. En este proyecto, sean $\{a_1, a_2, \ldots, a_n\}$, etcétera, conjuntos de números reales positivos.

(a) Se puede demostrar (por ejemplo, usando el teorema del valor medio que si a y b son positivos y $0 < \alpha < 1$ entonces

$$a^\alpha b^{1-\alpha} \le \alpha a + (1-\alpha) b$$

y que la igualdad es válida si y sólo si $a = b$. Suponiendo esto, sea $r > 1$ y considérese a s tal que

$$\frac{1}{r} + \frac{1}{s} = 1,$$

(de manera que $s > 1$ y $r + s = rs$). Demostrar que si A y B son positivos, entonces

$$AB \le \frac{A^r}{r} + \frac{B^s}{s},$$

y que la igualdad es válida si y sólo si $A^r = B^s$.

†JOSEPH-LOUIS LAGRANGE (1736-1813) nació en Turín, en donde se convirtió en profesor a los diecinueve años. Más tarde se fue a Berlín, por un periodo de veinte años, como sucesor de Euler, y después a París. Por lo que más se conoce es por su trabajo en el cálculo de variaciones y en mecánica analítica.

i083 (1628x2468x2 pcx)

(b) Sean $\{a_1, \ldots, a_n\}$ y $\{b_1, \ldots, b_n\}$ números reales positivos. Si $r, s > 1$ y $(1/r)+(1/s)=1$, probar la **desigualdad de Holder**†

$$\sum_{j=1}^{n} a_j b_j \leq \left\{\sum_{j=1}^{n} a_j^r\right\}^{1/r} \left\{\sum_{j=1}^{n} b_j^s\right\}^{1/s}.$$

(Sugerencia: Sean $A = \{\sum a_j^r\}^{1/r}$ y $B = \{\sum b_j^s\}^{1/s}$. Aplicar el inciso *(a)* a_j/A y b_j/B.)

(c) Usando la desigualdad de Holder, probar la **desigualdad de Minkowsky**‡

$$\left\{\sum_{j=1}^{n} (a_j + b_j)^r\right\}^{1/r} \leq \left\{\sum_{j=1}^{n} a_j^r\right\}^{1/r} + \left\{\sum_{j=1}^{n} b_j^r\right\}^{1/r}.$$

(Sugerencia: $(a+b)^r = (a+b)(a+b)^{r/s} = a(a+b)^{r/s} + b(a+b)^{r/s}$.)

(d) Usando la desigualdad de Holder, demostrar que

$$(1/n) \sum_{j=1}^{n} a_j \leq \left\{(1/n) \sum_{j=1}^{n} a_j^r\right\}^{1/r}.$$

(e) Si $a_1 \leq a_2$ y $b_1 \leq b_2$, entonces $(a_1 - a_2)(b_1 - b_2) \geq 0$ y por lo tanto

$$a_1 b_1 + a_2 b_2 \geq a_1 b_2 + a_2 b_1.$$

Probar que si $a_1 \leq a_2 \leq \cdots \leq a_n$ y $b_1 \leq b_2 \leq \cdots \leq b_n$, entonces

$$n \sum_{j=1}^{n} a_j b_j \geq \left\{\sum_{j=1}^{n} a_j\right\}\left\{\sum_{j=1}^{n} b_j\right\}.$$

(f) Suponga que $0 \leq a_1 \leq a_2 \leq \cdots \leq a_n 0 \leq b_1 \leq b_2 \leq \cdots \leq b_n$ y $r \geq 1$ Probar la **desigualdad de Chebyshev** †

$$\left\{(1/n) \sum_{j=1}^{n} a_j^r\right\}^{1/r} \left\{(1/n) \sum_{j=1}^{n} b_j^r\right\}^{1/r} \leq \left\{(1/n) \sum_{j=1}^{n} (a_j b_j)^r\right\}^{1/r}.$$

Demostrar que esta desigualdad se debe invertir cuando $\{a_j\}$ es creciente y $\{b_j\}$ es decreciente.

Sección 9 Conjuntos abiertos y cerrados

En el ánalisis real, muchas de las propiedades de mayor contenido dependen de ciertos conceptos topológicos. En las siguientes secciones se ha-

†OTTO HOLDER (1859-1937) estudió en Gottingen y dio clases en Leipzig. Trabajó tanto en álgebra como en análisis.

‡ HERMANN MINKOWSKI (1864-1909) fue profesor en Konigsberg y Gottingen. Por lo que más se le conoce es por su trabajo en conjuntos convexos y en la "geometría de números".

†PAFNUTI L. CHEBYSHEV (1821-1894) fue profesor en San Petesburgo. Hizo muchas contribuciones a las matemáticas, pero su trabajo más importante fue en teoría de números, probabilidad y teoría de la aproximación.

brán de introducir los conceptos básicos y establecer algunas de las propiedades topológicas más importantes del espacio $\boldsymbol{R}^p$. Estos resultados se usarán con frecuencia en los siguientes capítulos.

9.1 DEFINICION. Se dice que el conjunto G en $\boldsymbol{R}^p$ es **abierto en** $\boldsymbol{R}^p$ (o simplemente **abierto**) si para cada punto x en G existe un número real $r > 0$ tal que todo punto y en $\boldsymbol{R}^p$ que satisface $\|x - y\| < r$ también pertenece al conjunto G. (Fig. 9.1.)

Usando la definición 8.9 se puede reformular esta definición diciendo que un conjunto G es abierto si todo punto en G es el centro de alguna bola enteramente contenida en G.

9.2 EJEMPLOS. *(a)* El conjunto $\boldsymbol{R}^p$, en su totalidad, es abierto ya que se puede tomar $r = 1$ para cualquier x.

(b) El conjunto $G = \{x \in \boldsymbol{R} : 0 < x < 1\}$ es abierto en $\boldsymbol{R} = \boldsymbol{R}^1$. El conjunto $F = \{x \in \boldsymbol{R} : 0 \leq x \leq 1\}$ no es abierto en **R**. (¿por qué?)
qué?)

(c) Los conjuntos $G = \{(x, y) \in \boldsymbol{R}^2 : x^2 + y^2 < 1\}$ y $H = \{(x, y) : 0 < x^2 + y^2 < 1\}$ son abiertos pero el conjunto $F = \{(x, y) : x^2 + y^2 \leq 1\}$ no es abierto en $\boldsymbol{R}^2$. (¿por qué?)

(d) El conjunto $G = \{(x, y) \in \boldsymbol{R}^2 : 0 < x < 1,\ y = 0\}$ no es abierto en $\boldsymbol{R}^2$. Hacer una comparación con *(b)*. El conjunto $H = \{(x, y) \in \boldsymbol{R}^2 : 0 < y < 1\}$ es abierto pero el conjunto $K = \{(x, y) \in \boldsymbol{R}^2 : 0 \leq y < \}$ no es abierto en $\boldsymbol{R}^2$.

(e) El conjunto $G = \{(x, y, z) \in \boldsymbol{R}^3 : z > 0\}$ es abierto en $\boldsymbol{R}^3$ así como el conjunto $H = \{(x, y, z) \in \boldsymbol{R}^3 : x > 0,\ \ y > 0,\ \ z > 0\}$. Por otro lado, el conjunto $F = \{(x, y, z) \in \boldsymbol{R}^3 : x = y = z\}$ no es abierto.

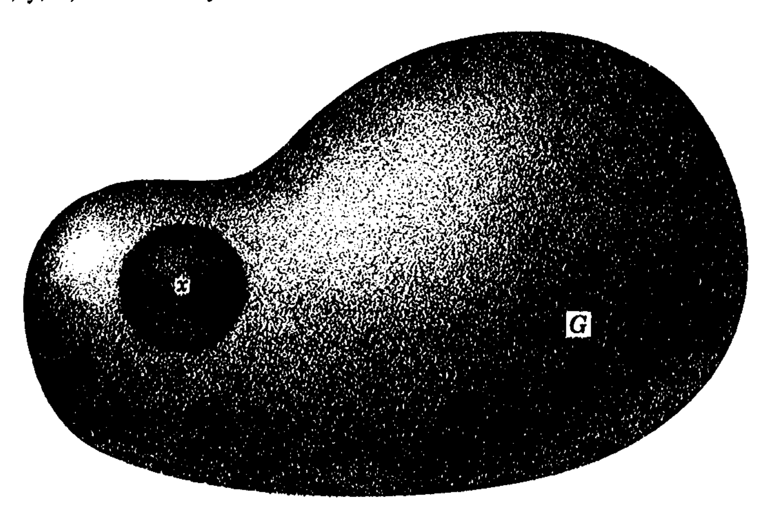


Figura 9.1. Un conjunto abierto.

www.elsolucionario.net

(f) El conjunto vacío $\emptyset$ es abierto en $\boldsymbol{R}^p$, ya que por no contener ningún punto el requisito de la definición 9.1, se satisface de manera trivial.

(g) Si B es la bola abierta con centro z y radio $a > 0$ y si $x \in B$, entonces la bola con centro x y radio $a - \|z - x\|$ está contenida en B. Por lo tanto, B es abierta en $\boldsymbol{R}^p$

En seguida se establecen las propiedades básicas de conjuntos abiertos en $\boldsymbol{R}^p$. En cursos de topología el siguiente resultado se resume diciendo que los conjuntos abiertos, como se establecen en la definición 9.1, forman una **topología** para $\boldsymbol{R}^p$

9.3 PROPIEDADES DE CONJUNTOS ABIERTOS. *(a) El conjunto vacío $\emptyset$ y la totalidad del espacio $\boldsymbol{R}^p$ son abiertos en $\boldsymbol{R}^p$.*

(b) La intersección de cualesquiera dos conjuntos abiertos es abierta en $\boldsymbol{R}^p$.

(c) La unión de cualquier colección de conjuntos abiertos es abierta en $\boldsymbol{R}^p$.

DEMOSTRACION. Ya se ha hecho referencia a la propiedad de abertura de los conjuntos $\emptyset$ y $\boldsymbol{R}^p$.

Para demostrar *(b)*, sean G_1, G_2 conjuntos abiertos y sea $G_3 = G_1 \cap G_2$. Para demostrar que G_3 es abierto, tome $x \in G_3$. Dado que x pertenece al conjunto abierto G_1, existe $r_1 > 0$ tal que si $\|x - z\| < r_1$, entonces $z \in G_1$. De manera análoga, existe $r_2 > 0$ tal que si $\|x - w\| < r_2$, entonces $w \in G_2$. Escogiendo a r_3 como el mínimo entre r_1 y r_2, se concluye que si $y \in \boldsymbol{R}^p$ es tal que $\|x - y\| < r_3$, entonces y pertenece tanto a G_1 como a G_2. Por lo tanto, tales elementos y pertenecen a $G_3 = G_1 \cap G_2$, demostrando que G_3 es abierto en $\boldsymbol{R}^p$.

Para demostrar *(c)*, sea $\{G_\alpha, G_\beta, \ldots\}$ una colección de conjuntos abiertos y sea G la unión de éstos. Para demostrar que G es abierto, tómese $x \in G$. De la definición de unión se infiere, que para algún conjunto, tómese por ejemplo G_λ, se tiene $x \in G_\lambda$. Dado que G_λ es abierto, existe una bola con centro x contenida enteramente en G_λ. Como $G_\lambda \subseteq G$, esta bola está contenida por completo en G, demostrando así que G es abierto en $\boldsymbol{R}^p$.

Por inducción se deduce de la propiedad *(b)* que la intersección de cualquier colecćción *finita* de conjuntos también es abierta en $\boldsymbol{R}^p$. El hecho de que la intersección de una colección infinita de conjuntos abiertos pueda no ser abierta se ve en el siguiente ejemplo:

$$G_n = \left\{x \in \boldsymbol{R} : -\frac{1}{n} < x < 1 + \frac{1}{n}\right\}, \qquad n \in \boldsymbol{N}. \tag{9.1}$$

La intersección de los conjuntos G_n es el conjunto $F = \{x \in \boldsymbol{R} : 0 \leq x \leq 1\}$, que no es abierto.

Conjuntos cerrados

En seguida se da el concepto importante de conjunto cerrado en $\boldsymbol{R}^p$.

9.4 DEFINICION. Se dice que un conjunto F en $\boldsymbol{R}^p$ es **cerrado** en $\boldsymbol{R}^p$ (o simplemente **cerrado**) cuando su complemento $\mathscr{C}(F) = \boldsymbol{R}^p \setminus F$ es abierto en $\boldsymbol{R}^p$.

9.5 EJEMPLOS. *(a)* La totalidad del conjunto $\boldsymbol{R}^p$ es cerrado en $\boldsymbol{R}^p$, ya que su complemento es el conjunto vacío y se ha visto en 9.2*(f)* que este conjunto es abierto en $\boldsymbol{R}^p$.

(b) El conjunto vacío $\emptyset$ es cerrado en $\boldsymbol{R}^p$, ya que su complemento en $\boldsymbol{R}^p$ es todo $\boldsymbol{R}^p$ y se ha visto en 9.2*(a)* que es abierto en $\boldsymbol{R}^p$.

(c) El conjunto $F = \{x \in \boldsymbol{R} : 0 \leq x \leq 1\}$ es cerrado en $\boldsymbol{R}$. Una manera de ver esto consiste en observar que el complemento de F en $\boldsymbol{R}$ es la unión de los dos conjuntos $\{x \in \boldsymbol{R} : x < 0\}$, $\{x \in \boldsymbol{R} : x > 1\}$, cada uno de los cuales es abierto. Análogamente, el conjunto $\{x \in \boldsymbol{R} : 0 \leq x\}$ es cerrado.

(d) El conjunto $F = \{(x, y) \in \boldsymbol{R}^2 : x^2 + y^2 \leq 1\}$ es cerrado, ya que su complemento en $\boldsymbol{R}^2$ es el conjunto

$$\{(x, y) \in \boldsymbol{R}^2 : x^2 + y^2 > 1\}.$$

y se puede ver que es abierto.

(e) El conjunto $H = \{(x, y, z) \in \boldsymbol{R}^3 : x \geq 0\}$ es cerrado en $\boldsymbol{R}^3$, así como el conjunto $F = \{(x, y, z) \in \boldsymbol{R}^3 : x = y = z\}$.

(f) La bola cerrada B con centro x en $\boldsymbol{R}^p$ y radio $r > 0$ es un conjunto cerrado de $\boldsymbol{R}^p$, ya que si $z \notin B$, entonces la bola abierta con centro z y radio $\|z - x\| - r$ está contenida en $\mathscr{C}(B)$. Por lo tanto, $\mathscr{C}(B)$ es abierto y B es cerrado en $\boldsymbol{R}^p$.

En el lenguaje ordinario, cuando se trata de puertas, ventanas y mentes, las palabras "abierto" y "cerrado" son antónimos. Sin embargo, cuando se aplican a subconjuntos de $\boldsymbol{R}^p$, estas palabras no son antónimos. Por ejemplo, ya se ha visto que los conjuntos $\emptyset$, $\boldsymbol{R}^p$ son *tanto* abiertos como cerrados en $\boldsymbol{R}^p$. (Probablemente el lector se sentirá conforme al saber que no hay más subconjuntos de $\boldsymbol{R}^p$ con ambas propiedades.) Además, hay muchos subconjuntos de $\boldsymbol{R}^p$ que no son *ni* abiertos *ni* cerrados; de hecho, la mayoría de los subconjuntos de $\boldsymbol{R}^p$ tienen esta propiedad neutral. Como un ejemplo sencillo se da el conjunto

(9.2) $$A = \{x \in \boldsymbol{R} : 0 \leq x < 1\}.$$

Este conjunto A no cumple con la propiedad de ser abierto en $\boldsymbol{R}$, ya que contiene al punto O. Análogamente, no cumple con la propiedad de ser cerrado en $\boldsymbol{R}$ porque su complemento en $\boldsymbol{R}$ es el conjunto $\{x \in \boldsymbol{R} : x < 0 \text{ ó } x \geq 1\}$, que no es abierto por contener al punto 1. El lector deberá construir otros ejemplos de conjuntos que no sean ni abiertos ni cerrados en $\boldsymbol{R}^p$.

En seguida se establecen las propiedades fundamentales de conjuntos cerrados. La demostración de este resultado se deduce directamente del teorema 9.3 usando las leyes de DeMorgan (teorema 1.8 y ejercicio 1.K).

9.6 PROPIEDADES DE CONJUNTOS CERRADOS. *(a) El conjunto vacío* $\emptyset$ *y todo el espacio* $\mathbf{R}^p$ *son cerrados en* $\mathbf{R}^p$.

(b) La unión de cualesquiera dos conjuntos cerrados es cerrada en $\mathbf{R}^p$.

(c) La intersección de cualquier colección de conjuntos cerrados es cerrada en $\mathbf{R}^p$.

Vecindades

Se introducirán ahora algunos conceptos adicionales de topología que resultarán útiles y harán posible la caracterización de conjuntos abiertos y cerrados en otros términos.

9.7 DEFINICION. *(a)* Si $x \in \mathbf{R}^p$, entonces cualquier conjunto que contenga un conjunto abierto conteniendo a x se llama una **vecindad** de x.

(b) A un punto $x \in \mathbf{R}^p$ se le llama un **punto interior** de un conjunto $A \subseteq \mathbf{R}^p$ cuando hay una vecindad de x enteramente contenida en A.

(c) A un punto $x \in \mathbf{R}^p$ se le llama un **punto frontera** de un conjunto $A \subseteq \mathbf{R}^p$ cuando toda vecindad de x contiene puntos en A y puntos en $\mathscr{C}(A)$.

(d) A un punto $x \in \mathbf{R}^p$ se le llama un **punto exterior** de un conjunto $A \subseteq \mathbf{R}^p$ cuando existe una vecindad de x enteramente contenida en $\mathscr{C}(A)$.

Se debe observar que dados $x \in \mathbf{R}^p$ y $A \subseteq \mathbf{R}^p$, hay tres posibilidades mutuamente excluyentes: (i) x es un punto interior de A, (ii) x es un punto frontera de A, o (iii) x es un punto exterior de A.

9.8 EJEMPLOS. *(a)* Un conjunto U es una vecindad de un punto x si y sólo si existe una bola con centro x enteramente contenida en U.

(b) Un punto x es un punto interior de A si y sólo si existe una bola con centro x enteramente contenida en A.

(c) Un punto x es un punto frontera de A si y sólo si para cada número natural n existen puntos $a_n \in A$ y $b_n \in \mathscr{C}(A)$ tales que $\|x - a_n\| < 1/n$ y $\|x - b_n\| < 1/n$.

(d) Todo punto del intervalo $(0, 1) \subseteq \mathbf{R}$ es un punto interior. Los puntos 0,1 son los puntos frontera de (0, 1).

(e) Sea $A = [0, 1] \subseteq \mathbf{R}$. Entonces, los puntos interiores de A son los puntos en el intervalo abierto (0, 1). Los puntos 0, 1 son los puntos frontera de A.

(f) Los puntos frontera de las bolas abiertas y cerradas con centro $x \in \mathbf{R}^p$ y radio $r > 0$, son los puntos de la esfera con centro x y radio r. (Véase definición 8.9)

En seguida se define a los conjuntos abiertos en términos de vecindades y de puntos interiores.

9.9 TEOREMA. *Si $B \subseteq \mathbf{R}^p$, entonces las siguientes afirmaciones son equivalentes:*

(a) B es abierto;

(b) todo punto de B es un punto interior de B;

(c) B es una vecindad de cada uno de sus puntos.

DEMOSTRACION. Si *(a)* se satisface y $x \in B$, entonces el conjunto abierto B es una vecindad de x y por lo tanto x es un punto interior de B.

Es trivial que *(b)* implica *(c)*.

Si se satisface *(c)* entonces para cada $x \in B$, hay un conjunto abierto $G_x \subseteq B$ con $x \in G_x$. Por lo tanto $B = \bigcup\{G_x : x \in B\}$, y se deduce del teorema 9.3(*c*) que B es abierto en $\mathbf{R}^p$.

A partir de lo que se ha demostrado, se deduce que un conjunto abierto no contiene a *ninguno* de sus puntos frontera. Los conjuntos cerrados son el extremo opuesto en lo que a esto respecta.

9.10 TEOREMA. *Un conjunto $F \subseteq \mathbf{R}^p$ es cerrado si y sólo si contiene a todos sus puntos frontera.*

DEMOSTRACION. Suponga que F es cerrado y que x es un punto frontera de F. Si $x \notin F$, entonces el conjunto abierto $\mathscr{C}(F)$ contiene a x y no contiene a ningún punto de F, contrario a la hipótesis de que x es un punto frontera de F. Por lo tanto, se debe tener $x \in F$.

A la inversa, suponga que F contiene a todos sus puntos frontera. Si $y \notin F$, entonces y no es ni un punto de F ni un punto frontera de F; por lo tanto, es un punto exterior. En consecuencia, existe una vecindad M de y enteramente contenida en $\mathscr{C}(F)$. Dado que esto es válido para toda $y \notin F$, se deduce que $\mathscr{C}(F)$ es abierto, de aquí que F sea cerrado en $\mathbf{R}^p$. Q.E.D.

Conjuntos abiertos en *R*

Se concluye esta sección haciendo una caracterización de la forma que tiene un subconjunto abierto arbitrario de $\mathbf{R}$.

9.11 TEOREMA. *Un subconjunto de R es abierto si y sólo si, es la union de colección contable de intervalos abiertos.*

DEMOSTRACION. Dado que un intervalo abierto es abierto (¿por qué?), de 9.3(*c*) se deduce que la unión de cualquier unión contable de intervalos abiertos es abierta.

A la inversa, sea $G \neq \emptyset$ un conjunto abierto en $\mathbf{R}$ y sea $\{r_n : n \in \mathbf{N}\}$ una enumeración de todos los puntos racionales en G. Para cada $n \in \mathbf{N}$ sea m_n el número natural más pequeño tal que el intervalo $J_n = (r_n - 1/m_n, r_n + 1/m_n)$ esté contenido integramente en G. Se deduce que

$$\bigcup_{n\in N} J_n \subseteq G.$$

Ahora, sea x un punto arbitrario en G y sea $m \in N$ tal que $(x - 2/m, x + 2/m) \subseteq G$. Del teorema 6.10 se deduce que existe un número racional y en $(x - 1/m, x + 1/m)$; de donde $y \in G$ y por lo tanto, $y = r_n$ para algún número natural n. Si x no pertenece a $J_n = (r_n - 1/m_n, r_n + 1/m_n)$, entonces se debe tener $1/m_n < 1/m$; pero como fácilmente se puede ver que

$$\left(r_n - \frac{1}{m}, r_n + \frac{1}{m}\right) \subseteq \left(x - \frac{2}{m}, x + \frac{2}{m}\right) \subseteq G,$$

esto contradice la elección de la m_n. Por lo tanto, se tiene $x \in J_n$ para este valor de n. Dado que $x \in G$ es arbitrario, se infiere que

$$G \subseteq \bigcup_{n\in N} J_n.$$

Por lo tanto, G es igual a esta unión.

Del teorema que se acaba de dar se deduce que un subconjunto de $\boldsymbol{R}$ es cerrado si y sólo si es la intersección de una colección contable de intervalos cerrados. (¿por qué?) *No* se deduce que la unión contable de intervalos cerrados deba ser cerrada, tampoco que todo conjunto cerrado tenga esta propiedad.

Una generalización de este resultado se da en el ejercicio 9.G.

Ejercicios

9.A. Justificar la afirmación que con respecto a los conjuntos G, F, se hizo en el ejemplo 9.2(b).

9.B. Justificar las afirmaciones hechas en el ejemplo 9.2(c).

9.C. Demostrar que la intersección de cualquier colección *finita* de conjuntos abiertos es abierta en $\boldsymbol{R}^p$. (*Sugerencia:* usar 9.3(b) e inducción.)

9.D. ¿Cuáles son los puntos interiores, frontera y exteriores del conjunto $(0,1)$ en R? Demostrar que no es abierto ni cerrado.

9.E. Dar un ejemplo en $\boldsymbol{R}^2$ que no sea ni abierto ni cerrado. Demostrar la afirmación.

9.F. Escribir los detalles de la demostración del teorema 9.6.

9.G. Demostrar que un subconjunto de $\boldsymbol{R}^p$ es abierto si y sólo si es la unión de una colección contable de bolas abiertas. (*Sugerencia:* el conjunto de todos los puntos en $\boldsymbol{R}^p$ para los cuales todas sus coordenadas son números racionales es contable.)

9.H. Todo subconjunto abierto de $\boldsymbol{R}^p$ es la unión de una colección contable de conjuntos cerrados.

9.I. Todo subconjunto cerrado de $\boldsymbol{R}^p$ es la intersección de una colección contable de conjuntos abiertos.

9.J. Si A es cualquier subconjunto de $\boldsymbol{R}^p$, denote mediante A^0 a la unión de todos los conjuntos abiertos que estén contenidos en A. Al conjunto A^0 se le llama el **interior** de A. Observe que A^0 es un conjunto abierto; demostrar que es el conjunto abierto más grande contenido en A. Demostrar que

$$A^0 \subseteq A, \qquad (A^0)^0 = A^0$$

$$(A \cap B)^0 = A^0 \cap B^0, \qquad (\boldsymbol{R}^p)^0 = \boldsymbol{R}^p.$$

Dar un ejemplo para demostrar que $(A \cup B)^0 = A^0 \cup B^0$ no es válido.

9.K. Demostrar que un punto pertenece a A^0 si y sólo si es un punto interior de A.

9.L. Si A es cualquier subconjunto de $\boldsymbol{R}^p$, denótese mediante A^- a la intersección de todos los conjuntos cerrados que contienen a A. Al conjunto A^- se le llama la **cerradura** de A. Observe que A^- es un conjunto cerrado. Demostrar que es el conjunto cerrado más pequeño que contiene a A. Demostrar que

$$A \subseteq A^-, \qquad (A^-)^- = A$$

$$(A \cup B)^- = A^- \cup B^-, \qquad \emptyset^- = \emptyset.$$

Dar un ejemplo para demostrar que $(A \cap B)^- = A^- \cap B^-$ no es válido.

9.M. Demostrar que un punto pertenece a A^- si y sólo si es un punto interior o bien un punto frontera de A.

9.N. Dar un ejemplo de un conjunto A en $\boldsymbol{R}^p$ tal que $A^0 = \emptyset$ y $A^- = \boldsymbol{R}^p$. ¿Puede ser contable dicho conjunto A?

9.O. Sean A y B subconjuntos de $\boldsymbol{R}$. El producto cartesiano $A \times B$ es abierto en $\boldsymbol{R}^2$ si y sólo si A y B son abiertos en $\boldsymbol{R}$.

9.P. Sean A y B subconjuntos de $\boldsymbol{R}$. El producto cartesiano $A \times B$ es cerrado en $\boldsymbol{R}^2$ si y sólo si A y B son cerrados en $\boldsymbol{R}$.

9.Q. Interpretar los conceptos introducidos en esta sección para el conjunto de Cantor F de la definición 7.4. Específicamente:

(a) Demostrar que F es cerrado en $\boldsymbol{R}$.

(b) No hay puntos interiores en F.

(c) No hay conjuntos abiertos, no vacíos contenidos en F.

(d) Todo punto de F es un punto frontera.

(e) El conjunto F no se puede expresar como la unión de una colección contable de intervalos cerrados.

(f) El complemento de F se puede expresar como la unión de una colección contable de intervalos abiertos.

Sección 10 Las celdas nidificadas y los teoremas de Bolzano-Weierstrass

En esta sección se darán a conocer dos resultados, muy importantes, que se usarán con frecuencia en capítulos posteriores. En cierto sentido, se pueden

considerar como la propiedad de completación para $\mathbf{R}^p$, cuando $p > 1$. En la sección 7 se ha visto que si $a \leq b$, entonces la celda abierta en $\boldsymbol{R}$, designada con (a, b), es el conjunto definido por

$$(a, b) = \{x \in \boldsymbol{R} : a < x < b\}.$$

Se puede ver fácilmente que dicho conjunto es abierto en $\boldsymbol{R}$. De manera análoga, la celda cerrada $[a, b]$ en $\boldsymbol{R}$ es el conjunto

$$[a, b] = \{x \in \boldsymbol{R} : a \leq x \leq b\},$$

que es cerrado en $\boldsymbol{R}$. El producto cartesiano de dos intervalos, por lo general, se llama un **rectángulo** y el producto cartesiano a tres intervalos a menudo se llama un **paralelepípedo.** Por comodidad se empleará el término "celda" haciendo caso omiso de la dimensión del espacio.

10.1 DEFINICION. Una **celda abierta** J en $\boldsymbol{R}^p$ es el producto cartesiano de p celdas abiertas de números reales. Por lo tanto, J es de la forma

$$J = \{x = (x_1, \ldots, x_p) \in \boldsymbol{R}^p : a_i < x_i < b_i, \quad \text{para} \quad i = 1, 2, \ldots, p\}.$$

Análogamente, una **celda cerrada** I en $\mathbf{R}^p$ es el producto cartesiano de p celdas cerradas de números reales. Por lo tanto, I es de la forma

$$I = \{x = (x_1, \ldots, x_p) \in \boldsymbol{R}^p : a_i \leq x_i \leq b_i, \quad \text{para} \quad i = 1, 2, \ldots, p\}.$$

Un subconjunto de $\boldsymbol{R}^p$ es **acotado** si está contenido en alguna celda.

Como ejercicio, demostrar que una celda abierta en $\boldsymbol{R}^p$ es un conjunto abierto y que una celda cerrada es un conjunto cerrado. También demostrar que un subconjunto de $\boldsymbol{R}^p$ es acotado si y sólo si está contenido en alguna bola. Se observará que esta terminología para conjuntos acotados es congruente con la que se dio a conocer en la sección 6 para el caso $p = 1$.

El lector recordará que en la sección 7 la propiedad del supremo para el sistema de números reales implica que toda sucesión nidificada de celdas no vacías cerradas en $\boldsymbol{R}$ tiene un punto en común. En seguida se demostrará que esta propiedad sigue siendo válida para el espacio $\boldsymbol{R}^p$.

10.2 TEOREMA DE CELDAS NIDIFICADAS. *Sea (I_k) una sucesión de celdas no vacías cerradas en $\mathbf{R}^p$ nidificada en el sentido de que $I_1 \supseteq I_2 \supseteq \cdots \supseteq I_k \supseteq \cdots$. Entonces, existe un punto en $\mathbf{R}^p$ que pertenece a todas las celdas.*

DEMOSTRACION. Suponga que I_k es la celda

$$I_k = \{(x_1, \ldots, x_p) : a_{k1} \leq x_1 \leq b_{k1}, \ldots, a_{kp} \leq x_p \leq b_{kp}\}.$$

Es fácil ver que las celdas $[a_{k1}, b_{k1}], k \in \boldsymbol{N}$, forman una sucesión nidificada de celdas cerradas no vacías de números reales y por la completación del sistema de números reales $\boldsymbol{R}$, hay un número real y_1 que pertenece a todas estas celdas. Aplicando dicho argumento a cada coordenada, se obtiene un punto $y = (y_1, \ldots, y_p)$ de $\boldsymbol{R}^p$ tal que si j satisface $j = 1, 2, \ldots, p$, entonces y_j pertenece a todas las celdas $\{[a_{kj}, b_{kj}] : k \in \boldsymbol{N}\}$. Por lo tanto, el punto y pertenece a todas las celdas (I_k). Q.E.D.

Puntos de acumulación y Bolzano-Weierstrass

10.3 DEFINICION. Un punto $x \in \boldsymbol{R}^p$ es un **punto de acumulación** (o **punto límite**) de un subconjunto $A \subseteq \boldsymbol{R}^p$ cuando toda vecindad de x contiene cuando menos un punto de A distinto de x.

En seguida se dan algunos ejemplos.

10.4 EJEMPLOS. *(a)* Un punto $x \in \boldsymbol{R}^p$ es un punto de acumulación de A si y sólo si para todo número natural n existe un elemento $a_n \in A$ tal que $0 < \|x - a_n\| < 1/n$.

(b) Si un punto frontera de un conjunto no pertenece al conjunto, entonces es un punto de acumulación del conjunto.

(c) Todo punto del intervalo unitario $\boldsymbol{I}$ de $\boldsymbol{R}$ es un punto de acumulación de $\boldsymbol{I}$.

(d) Sea $A = (0, 1)$, entonces todo punto de A es punto interior así como punto de acumulación de A. Los puntos 0,1 son puntos de acumulación (y no puntos interiores) de A.

(e) Sea xxxxxxxx el conjunto de todos los números racionales del intervalo unitario. Todo punto I es un punto de acumulación de B en $\mathbf{R}$, pero no hay puntos interiores de B.

(f) Un subconjunto finito de xxx no tiene puntos de acumulación. (¿por qué?)

(g) El conjunto infinito de enteros xxxxxi no tiene puntos de acumulación. (¿por qué?)

10.5 TEOREMA. *Un conjunto* xxxxxxi *es cerrado si y sólo si contiene a todos sus puntos de acumulación.*

acumulación de F. Si $x \notin F$, el conjunto abierto $\mathscr{C}(F)$ es una vecindad de x, entonces debe contener cuando menos un punto de F. Pero esto no es posible; por lo tanto, se concluye que $x \in F$.

A la inversa, si F contiene a todos sus puntos de acumulación, se habrá de probar que $\mathscr{C}(F)$ es abierto, ya que si $y \in \mathscr{C}(F)$, entonces y no es punto de acumulación de F. Por lo tanto, existe una vecindad V_y de y tal que $F \cap V_y = \emptyset$. En consecuencia $V_y \subseteq \mathscr{C}(F)$. Dado que esto es válido para toda $y \in \mathscr{C}(F)$, se deduce que $\mathscr{C}(F)$ es abierto en $\boldsymbol{R}^p$. Q.E.D.

El siguiente resultado es uno de los más importantes en este libro. Es de gran importancia y se usará con frecuencia. Se debe observar que al omitir alguna hipótesis la conclusión puede fallar (ver ejemplos 10.4(f,g)).

10.6. TEOREMA DE BOLZANO-WEIERSTRASS†. *Todo subconjunto infinito acotado de* $\boldsymbol{R}^p$ *tiene un punto de acumulación.*

DEMOSTRACION. Si B es un conjunto acotado con un número infinito de elementos, suponga que I_1 es una celda cerrada que contiene a B. Se divide I_1 en 2^p celdas cerradas bisecando cada uno de sus lados. Dado que I_1 contiene una infinidad de puntos de B, al menos una de las partes obtenidas en esta subdivisión también contendrá una infinidad de puntos de B. (Puesto que si cada una de las 2^p partes tuviera sólo un número finito de puntos del conjunto B, entonces B sería un conjunto finito, contrario a la hipótesis.) Sea I_2 una de estas partes de la subdivisión de I_1 que contiene una infinidad de elementos de B. Divida ahora I_2 en 2^p celdas cerradas bisecando cada uno de sus lados. De nuevo, una de estas subceldas de I_2 debe contener un número infinito de puntos de B, ya que, de no ser as así, I_2 podría contener solamente un número finito, contrario a su construcción. Sea I_3 una subcelda de I_2 que contenga una infinidad de puntos de B. Si se continúa este procedimiento, se obtiene una sucesión nidificada (I_k) de celdas cerradas no-vacías de $\boldsymbol{R}^p$. Según el teorema de celdas nidificadas, existe un punto y que pertenece a todas las celdas I_k, $k = 1, 2, \ldots$. En seguida se probará que y es un punto de acumulación de B y con esto se completará la demostración de la afirmación.

En primer lugar, observe que si $I_1 = [a_1, b_1] \times \cdots \times [a_p, b_p]$ con $a_k < b_k$, y si $l(I_1) = \sup\{b_1 - a_1, \ldots, b_p - a_p\}$, entonces $l(I_1) > 0$ es la longitud del lado mayor de I_1. De acuerdo con la construcción anterior de la sucesión (I_k), se tiene

$$0 < l(I_k) = \frac{1}{2^{k-1}}\, l(I_1)$$

para $k \in \boldsymbol{N}$. Suponga que V es cualquier vecindad del punto en común y suponga que todos los puntos z en $\boldsymbol{R}^p$ tales que $\|y - z\| < r$ pertenecen a V. Se va a elegir k de tal magnitud que $I_k \subseteq V$; esta elección es posible ya que si w es cualquier otro punto de I_k, entonces del teorema 8.10 se deduce que

†BERNARD BOLZANO (1781-1897) fue profesor de filosofía de la religión en Praga, pero hacía reflexiones profundas acerca de las matemáticas. Igual que Cauchy, fue un pionero al introducir un nivel más alto en el rigor del análisis matemático. Sus tratados de las paradojas del infinito aparecieron después de su muerte.

KARL WEIERSTRASS (1815-1897), durante muchos años fue profesor en Berlín y ejerció una profunda influencia en el desarrollo del análisis. Insistiendo siempre en las demostraciones rigurosas desarrolló, pero no publicó, una introducción al sistema de números reales. También hizo importantes aportaciones al análisis real y complejo, las ecuaciones diferenciales y el cálculo de variaciones.

$$\|y - w\| \le \sqrt{p}\, l(I_k) = \frac{\sqrt{p}}{2^{k-1}}\, l(I_1).$$

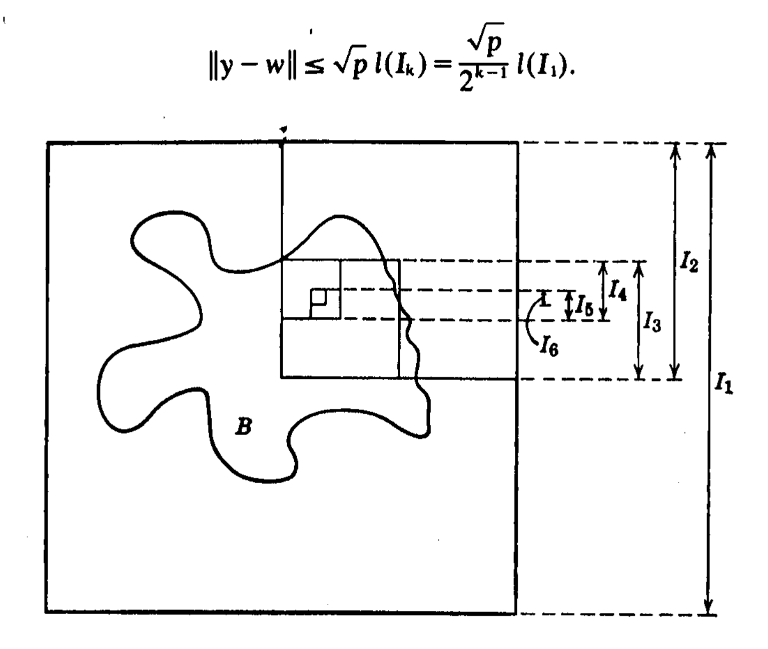


Figura 10.1

Del corolario 6.7 se deduce que si k es suficientemente grande, entonces

$$\frac{\sqrt{p}}{2^{k-1}}\, l(I_1) < r.$$

Para tal valor de k se tiene $I_k \subseteq V$. Dado que I_k contiene una infinidad de elementos de B, se infiere que V contiene cuando menos un elemento de B distinto de y. Por lo tanto, y es un punto de acumulación de B.

Ejercicios

10.A. Sean $I_n \subseteq \mathbf{R}^p$ las celdas abiertas dadas por $I_n = (0, 1/n) \times \cdots \times (0, 1/n)$. Demostrar que estas celdas son nidificadas pero que no contienen ningún punto en común.

10.B. Sean $J_n \subseteq \mathbf{R}^p$ los intervalor cerrados dados por $J_n = [n, +\infty) \times \cdots \times [n, +\infty)$. Demostrar que estos intervalos son nidificados pero que no contienen ningún punto en común.

10.C. Un punto x es un punto de acumulación de un conjunto $A \subseteq \mathbf{R}^p$ si y sólo si toda vecindad de x contiene una infinidad de puntos de A.

10.D. Sea $A = \{1/n : n \in \mathbf{N}\}$. Demostrar que todo punto de A es un punto frontera en $\mathbf{R}$ pero que 0 es el único punto de acumulación de A en $\mathbf{R}$.

10.E. Sean A, B subconuntos de $\mathbf{R}^p$ y sea x un punto de acumulación de $A \cap B$ en $\mathbf{R}^p$. Demostrar que x es un punto de acumulación tanto de A como de B.

10.F. Sean A, B subconjuntos de $\mathbf{R}^p$ y sea x un punto de acumulación de $A \cup B$ en $\mathbf{R}^p$. Demostrar que x es punto de acumulación de A o de B.

10.G. Demostrar que todo punto del conjunto de Cantor F es un punto de acumulación tanto de F como de $\mathscr{C}(F)$.

10.H. Si A es cualquier subconjunto de $\mathbf{R}^p$, existe un subconjunto contable C de A tal que si $x \in A$ y $\varepsilon > 0$, entonces, hay un elemento $z \in C$ tal que $\|x - z\| < \varepsilon$. Por lo tanto, todo elemento de A está en C o bien es un punto de acumulación de C.

Proyectos

10.α. Sea M un conjunto y sea d una métrica en M como se definió en el ejercicio 8.5. Reexaminar las definiciones y los teoremas de las secciones 9 y 20 para determinar cuáles se siguen cumpliendo en conjuntos que tienen una métrica. Por ejemplo, se verá que son válidos los conceptos de conjunto abierto, cerrado y acotado. Sin embargo, el teorema de Bolzano-Weierstrass falla para ciertos M y d. Siempre que sea posible, demostrar que el teorema se extiende, o bien dar un contraejemplo para demostrar que puede fallar.

10.β. Sea $\mathscr{T}$ una familia de subconjuntos de un conjunto X tal que (i) contiene a $\emptyset$ y X, (ii) contiene a la intersección de cualquier familia finita de conjuntos en $\mathscr{T}$, y (iii) contiene a la unión de cualquier familia de conjuntos en $\mathscr{T}$. Se dice que $\mathscr{T}$ es una topología para X y se hace referencia a los conjuntos en $\mathscr{T}$ como los conjuntos abiertos. Reexaminar las definiciones y los teoremas de las secciones 9 y 10 tratando de determinar cuáles persisten en conjuntos X que tienen una topología $\mathscr{T}$.

Seccción 11 El teorema de Heine-Borel

El teorema de celdas nidificadas 10.2 y el teorema de Bolzano-Weierstrass 10.6 están íntimamente relacionados al concepto tan importante de compacidad que se analiza en esta sección. Aunque es posible obtener la mayoría de resultados de las siguientes secciones sin conocer el teorema de Heine-Borel, no se puede avanzar mucho más dentro del análisis sin requerir de este teorema, de modo que no conviene tratar de evadir la exposición de este resultado tan profundo.

11.1 DEFINICION. Se dice que un conjunto K es **compacto** si siempre que está contenido en la unión de una colección $\mathscr{G} = \{G_\alpha\}$ de conjuntos abiertos también esta contenido en la unión de algún número *finito* de conjuntos en $\mathscr{G}$.

Una colección $\mathscr{G}$ de conjuntos abiertos cuya unión contiene a K con frecuencia se llama una **cubierta** de K. De modo que el requisito para que K sea compacto es que toda cubierta $\mathscr{G}$ de K se pueda sustituir por una cubierta *finita* de K, usando únicamente conjuntos de $\mathscr{G}$. Obsérvese que para aplicar esta definición al demostrar que un conjunto K es compacto, se debe examinar una colección arbitraria de conjuntos abiertos cuya unión contenga a K y demostrar que K está contenido en la unión de alguna subcolección finita de cada una de estas colecciones. Por otro lado, para demostrar que un conjunto

H no es compacto, basta con exhibir sólo *una* cubierta que no se pueda reemplazar por una subcolección finita que siga cubriendo a H.

11.2 EJEMPLOS. *(a)* Sea $K = \{x_1, x_2, \ldots, x_m\}$ un subconjunto finito de $\boldsymbol{R}^p$. Es claro que si $\mathcal{G} = \{G_\alpha\}$ es una colección de conjuntos abiertos en $\boldsymbol{R}^p$, y si todo punto de K pertenece a algún subconjunto de $\mathcal{G}$., entonces cuando más m subconjuntos de $\mathcal{G}$, que se seleccionen cuidadosamente, también tendrán la propiedad de que su unión contenga a K. Por lo tanto, K es un subconjunto compacto de $\boldsymbol{R}^p$.

(b) Considere al subconjunto $H = \{x \in \boldsymbol{R} : x \geq 0\}$. en $\boldsymbol{R}$. Sea $G_n = (-1, n)$, $n \in \boldsymbol{N}$, de tal manera que $\mathcal{G} = \{G_n : n \in \boldsymbol{N}\}$ sea una colección de subconjuntos abiertos de $\mathbf{R}$ cuya unión contenga a H. Si $\{G_{n_1}, G_{n_2}, \ldots, G_{n_k}\}$ es una subcolección finita de $\mathcal{G}$, sea $M = \sup\{n_1, n_2, \ldots, n_k\}$ de tal manera que $G_{n_j} \subseteq G_M$, para $j = 1, 2, \ldots, k$. Se deduce que G_M es la unión de $\{G_{n_1}, G_{n_2}, \ldots, G_{n_k}\}$. Sin embargo, el número real M no pertenece a G_M y por lo tanto no pertenece a

$$\bigcup_{j=1}^{k} G_{n_j}.$$

En consecuencia, ninguna unión finita de los conjuntos $\mathcal{G}$ puede contener a H y H no es compacto.

(c) Sea $H = (0, 1)$ en $\boldsymbol{R}$. Si $G_n = (1/n, 1 - 1/n)$ para $n > 2$, entonces la coleccion $\mathcal{G} = \{G_n : n > 2\}$ de conjuntos abiertos es una cubierta de H. Si $\{G_{n_1}, \ldots, G_{n_k}\}$ es una subcolección finita de $\mathcal{G}$, sea $M = \sup\{n_1, \ldots, n_k\}$ de tal manera que $G_{n_j} \subseteq G_M$ para $j = 1, 2, \ldots, k$. Se infiere que G_M es la unión de los conjuntos $\{G_{n_1}, \ldots, G_{n_k}\}$. Sin embargo, el número real $1/M$ pertenece a H pero no pertenece a G_M. Por lo tanto, ninguna subcolección finita de $\mathcal{G}$ puede formar una cubierta de H, de modo que H no es compacto.

(d) Considere el conjunto $\boldsymbol{I} = [0, 1]$; se demostrará que $\boldsymbol{I}$ es compacto. Sea $\mathcal{G} = \{G_\alpha\}$ una colección de subconjuntos abiertos de $\boldsymbol{R}$ cuya unión contiene a $\boldsymbol{I}$. El número real $x = 0$ pertenece a algún conjunto abierto de la colección $\mathcal{G}$ al igual que los números x que satisfacen $0 \leq x < \varepsilon$, para alguna $\varepsilon > 0$. Sea x^* el supremo de aquellos puntos x en $\boldsymbol{I}$ tales que la celda $[0, x]$ esté contenida en la unión de un número finito de conjuntos en $\mathcal{G}$. Dado que x^* pertenece a $\boldsymbol{I}$, se deduce que x^* es un elemento de algún conjunto abierto en $\mathcal{G}$. Por lo tanto, para alguna $\varepsilon > 0$, la celda $[x^* - \varepsilon, x^* + \varepsilon]$ está contenida en un conjunto G_0 en la colección $\mathcal{G}$. Pero (por la definición de x^*) la celda $[0, x^* - \varepsilon]$ está contenida en la unión de un número finito de conjuntos en $\mathcal{G}$. Por lo tanto, sumando el conjunto G_0 al número finito de conjuntos necesarios para cubrir $[0, x^* - \varepsilon]$, se deduce que el conjunto $[0, x^* + \varepsilon]$ está contenido en la unión de un número finito de conjuntos en $\mathcal{G}$. Esto da una contradicción a menos de que $x^* = 1$.

Por lo general, no es fácil demostrar que un conjunto es compacto usando únicamente la definición. En seguida se ofrece un teorema notable e

importante que caracteriza por completo a subconjuntos compactos de $\boldsymbol{R}^p$. De hecho, parte de la importancia del teorema de Heine-Borel† se debe a la sencillez de las condiciones de compacidad en $\boldsymbol{R}^p$.

11.3 TEOREMA DE HEINE-BOREL. *Un subconjunto de* $\boldsymbol{R}^p$ *es compacto si y sólo si es cerrado y acotado.*

DEMOSTRACION. Primero se demuestra que si K es compacto en $\boldsymbol{R}^p$, entonces K es cerrado. Sea x un elemento de $\mathscr{C}(K)$ y para cada número natural m sea G_m el conjunto definido por

$$G_m = \{y \in \boldsymbol{R}^p : \|y - x\| > 1/m\}.$$

Fácilmente se puede ver que cada conjunto G_m, $m \in \boldsymbol{N}$, es abierto en $\boldsymbol{R}^p$. Además, la unión de todos los conjuntos G_m, $m \in \boldsymbol{N}$, consta de todos los puntos de $\boldsymbol{R}^p$ excepto x. Dado que $x \notin K$, cada punto de K pertenece a algún conjunto G_m. Debido a la compacidad de K, se infiere que existe un número natural M tal que K está contenido en la unión de los conjuntos

$$G_1, G_2, \ldots, G_M.$$

Dado que los conjuntos G_m incrementan con m, K está contenido en G_M. De donde la vecindad $\{z \in \boldsymbol{R}^p : \|z - x\| < 1/M\}$ no intercepta a K, demostrando que $\mathscr{C}(K)$ es abierto. Por lo tanto, K es cerrado en $\boldsymbol{R}^p$. (Fig. 11.1 en donde las bolas cerradas complementarias a G_m están dibujadas.)

En seguida se demuestra que si K es compacto en $\boldsymbol{R}^p$, entonces K está acotado (es decir, K está contenido en algún conjunto $\{x \in \boldsymbol{R}^p : \|x\| < r\}$ para r lo suficientemente grande). De hecho, para cada número natural m, sea H_m el conjunto abierto definido por

$$H_m = \{x \in \boldsymbol{R}^p : \|x\| < m\}.$$

Todo el espacio $\boldsymbol{R}^p$, y por lo tanto K, está contenido en la unión de los conjuntos crecientes H_m, $m \in \boldsymbol{N}$. Dado que K es compacto, existe un número natural M tal que $K \subseteq H_M$. esto prueba que K está acotado.

Para completar la demostración de este teorema se necesita probar que si K es un conjunto cerrado y acotado contenido en la unión de una colección $\mathscr{G} = \{G_\alpha\}$ de conjuntos abiertos en $\boldsymbol{R}^p$, entonces está contenido en la unión de algún número finito de conjuntos en $\mathscr{G}$. Dado que el conjunto K está acotado,

†EDUARD HEINE (1821-1881) estudió en Berlín bajo la tutela de Weierstrass y más tarde dio clases en Bonn y Halle. En 1872 demostró que una función continua en un intervalo cerrado es uniformemente continua.

(F.E.J.) EMILE BOREL (1871-1956), alumno de Hermite, fue profesor en París y uno de los matemáticos más influyentes de su época. Hizo numerosas y profundas aportaciones al análisis y a la probabilidad. En 1895 demostró que si una colección contable de intervalos abiertos cubre a un intervalo cerrado entonces tiene una subcubierta finita.

i098 (1544x2467x2 pcx)

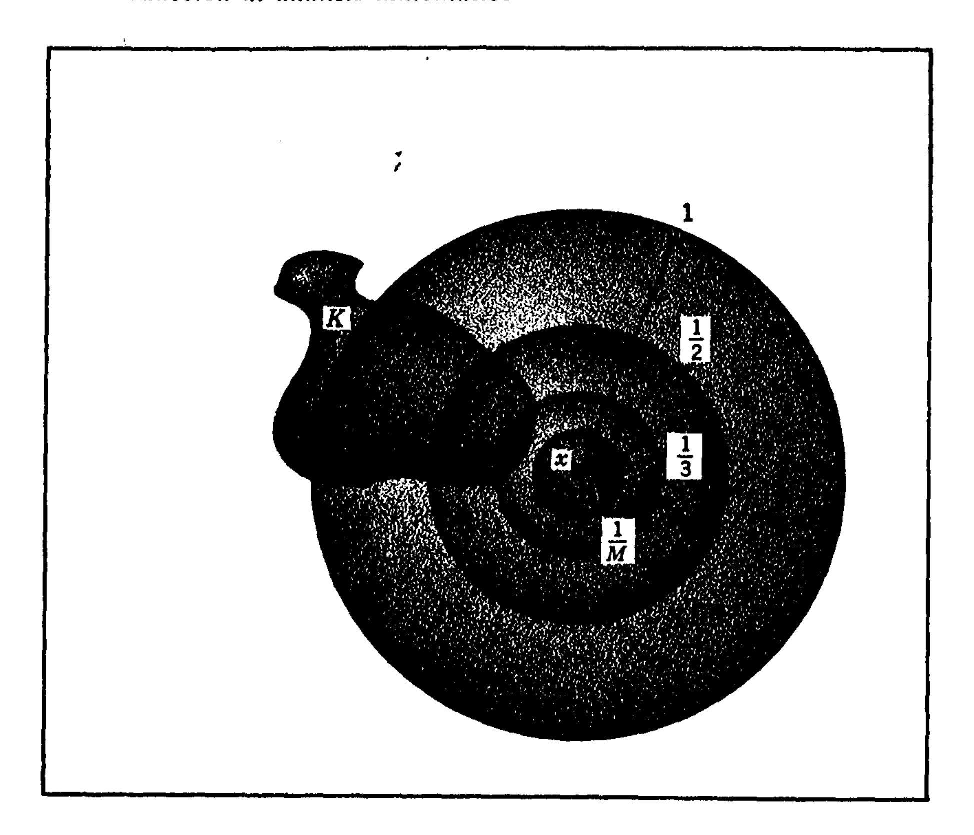


Figura 11.1. Un conjunto compacto es cerrado.

se puede encerrar en una celda cerrada I_1 en $\boldsymbol{R}^p$. Por ejemplo, se puede tomar $I_1 = \{(x_1, \ldots, x_p) : |x_k| \leq r,\ k = 1, \ldots, p\}$ para $r > 0$. lo suficientemente grande. Con el objeto de obtener una contradicción se habrá de suponer que K no está contenido en la unión de cualquier número finito de los conjuntos en $\mathcal{G}$. Por lo tanto, al menos una de las 2^p celdas cerradas obtenidas al bisecar los lados de I_1 contiene puntos de K y es tal que la parte de K que está en ella no está contenida en la unión de cualquier número finito de los conjuntos en $\mathcal{G}$. (puesto que si cada una de las 2^p partes de K estuvieran contenidas en la unión de un número finito de conjuntos en $\mathcal{G}$, entonces K estaría contenido en la unión de un número finito de conjuntos en $\mathcal{G}$, contrario a la hipótesis). Sea I_2 cualquiera de las subceldas de esta subdivisión de I_1 tal que el conjunto no vacío $K \cap I_2$ no esté contenido en la unión de cualquier número finito de conjuntos en $\mathcal{G}$. Se continúa este proceso bisecando los lados de I_2 para obtener 2^p subceldas cerradas de I_2 y se considera a I_3 como una de estas subseldas tal que el conjunto no vacío $K \cap I_3$ no está contenido en la unión de un número infinito de conjuntos en $\mathcal{G}$, y así sucesivamente.

De esta manera se obtiene una sucesión nidificada (I_n) de celdas no vacías. (Fig. 11.2.) De acuerdo con el teorema de celdas nidificadas, hay un punto y común para las I_n. Dado que cada I_n contiene puntos en K, el ele-

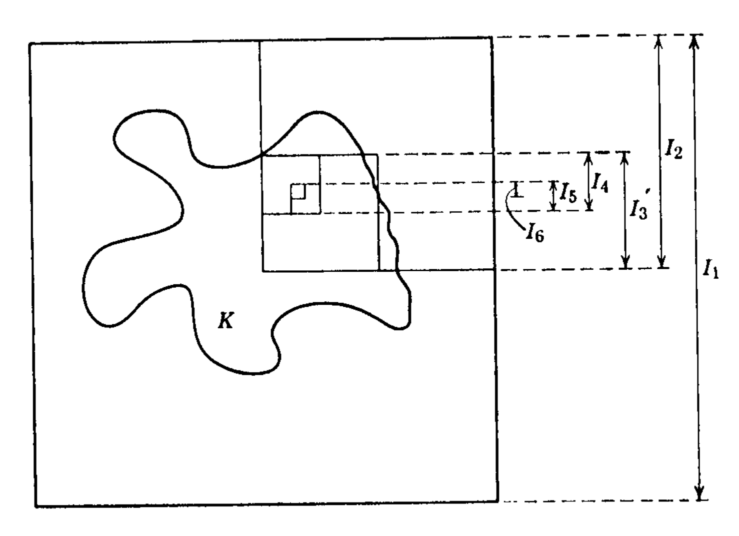


Figura 11.2

mento en común y es un punto de acumulación de K. Como K es cerrado, y pertenece a K y está contenido en algún conjunto abierto G_λ en $\mathcal{G}$. Por lo tanto, existe un número $\varepsilon > 0$ tal que todos los puntos w con $\|y-w\| < \varepsilon$ pertenece a G_λ. Por otro lado, las celdas I_k, $k \geq 2$, se obtienen mediante la visección sucesiva de los lados de la celda $I_1 = \{(x_1 \ldots, x_p) : |x_j| \leq r\}$ de tal manera que la magnitud del lado de I_k sea $r/2^{k-2}$. Del teorema 8.10 se deduce que si $w \in I_k$, entonces $\|y-w\| \leq r\sqrt{p}/2^{k-1}$. Por lo tanto, si K se escoge de tal magnitud que $r\sqrt{p}/2^{k-1} < \varepsilon$, entonces todos los puntos en I_k están contenidos en el único conjunto G_λ. Pero esto contradice la construcción de I_k como un conjunto tal que $K \cap I_k$ no esté contenido en la unión de un número finito de conjuntos en $\mathcal{G}$. Dicha contradicción muestra que es insostenible la suposición de que el conjunto cerrado acotado K requiere de un número infinito de conjuntos en $\mathcal{G}$ para encerrarlo.

Algunas aplicaciones

Como consecuencia del teorema de Heine-Borel se obtiene el siguiente resultado que se debe a Cantor. Es un reforzamiento del teorema de celdas nidificadas ya que en este caso se toma en cuenta a los conjuntos cerrados en general y no solamente a las celdas cerradas.

11.4 TEOREMA DE INTERSECCION DE CANTOR. *Sea F_1 un subconjunto no vacío, cerrado y acotado de $\mathbf{R}^p$ y sea*

$$F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$$

una sucesión de conjuntos cerrados no vacíos. Entonces, existe un punto que pertenece a todos los conjuntos $\{F_k : k \in \mathbf{N}\}$.

DEMOSTRACION. Dado que F_1 es cerrado y acotado, del teorema de Heine-Borel se deduce que es compacto. Para cada $k \in \mathbf{N}$, sea G_k el complemento de F_k en $\mathbf{R}^p$ como se ha supuesto que F_k es cerrado, G_k es abierto en $\mathbf{R}^p$ si, contrario al teorema, no hay ningún punto que pertenezca a todos los conjuntos F_k, $k \in \mathbf{N}$, entonces la unión de los conjuntos G_k, $k \in \mathbf{N}$, contiene al conjunto compacto F_1. Por lo tanto, el conjunto F_1 está contenido en la unión de un número finito de los conjuntos G_k; digamos en $G_1, G_2, \ldots, G_K$. Dado que los conjuntos G_k son crecientes, se tiene $G_1 \cup \cdots \cup G_K = G_K$. como $F_1 \subseteq G_K$, se deduce que $F_1 \cap F_K = \emptyset$. Por hipótesis $F_1 \supseteq F_K$, de modo que $F_1 \cap F_K = F_K$. El supuesto nos lleva a la conclusión de que $F_K = \emptyset$, que contradice la hipótesis y ratifica el teorema.

Q.E.D.

11.5 TEOREMA DE COBERTURA DE LEBESGUE. *Suponga que $\mathcal{G} = \{G_\alpha\}$ es una cubierta de un subconjunto compacto K de $\mathbf{R}^p$. Existe un número λ estrictamente positivo tal que si x, y pertenecen a K y $\|x - y\| < \lambda$, entonces hay un conjunto en $\mathcal{G}$ que contiene tanto a x como a y.*

DEMOSTRACION. Para cada punto u en K hay un conjunto abierto $G_{\alpha(u)}$ en $\mathcal{G}$ que contiene a u. Sea $\delta(u) > 0$ tal que si $\|v - u\| < 2\delta(u)$, entonces v pertenece a $G_{\alpha(u)}$. Considere el conjunto abierto $S(u) = \{v \in \mathbf{R}^p : \|v - u\| < \delta(u)\}$ y la colección $\mathcal{S} = \{S(u) : u \in K\}$ de conjuntos abiertos. Dado que $\mathcal{S}$ es una cubierta del conjunto compacto K, entonces K está contenido en la unión de un número finito de conjuntos $\mathcal{S}$, digamos en $S(u_1), \ldots, S(u_n)$. Se define λ como el número real estrictamente positivo

$$\lambda = \inf\{\delta(u_1), \ldots, \delta(u_n)\}.$$

Si x, y pertenecen a K y $\|x - y\| < \lambda$, entonces x pertenece a $S(u_j)$ para alguna j con $1 \leq j \leq n$, de manera que $\|x - u_j\| < \delta(u_j)$. Dado que $\|x - y\| < \lambda$, se tiene $\|y - u_j\| \leq \|y - x\| + \|x - u_j\| < 2\delta(u_j)$. De acuerdo con la definición de $\delta(u_j)$, se deduce que tanto x como y pertenecen al conjunto $G_{\alpha(u_j)}$.

Se hace la observación de que a un número positivo λ con la propiedad dada en el teorema, algunas veces se le llama un **número de Lebesque**† de la cubierta $\mathcal{G}$.

A pesar de que en secciones posteriores se estarán usando argumentos basados en compacidad, resulta apropiado intercalar aquí dos resultados que son intuitivamente claros pero cuya demostración parece requerir del uso de algún tipo de argumento acerca de la compacidad.

11.6 TEOREMA DEL PUNTO MAS PROXIMO. *Sea F en subconjunto cerrado no vació de $\mathbf{R}^p$ y sea x un punto fuera de F. Entonces existe*

†HENRI LEBESGUE (1875-1941), Por lo que es más conocido es por su trabajo de exploración en la teoría moderna de la integral que lleva su nombre y que es básica en el análisis actual.

cuando menos un punto y perteneciente a F tal que $\|z-x\| \geq \|y-x\|$ para toda $z \in F$.

DEMOSTRACION. Dado que F es cerrado y $x \notin F$, (cf. ejercicio 11.H) la distancia de x a F, que se define como $d = \inf\{\|x-z\| : z \in F\}$ satisface $d > 0$. Sea $F_k = \{z \in F : \|x-z\| \leq d + 1/k\}$ para $k \in \mathbf{N}$. De acuerdo con el ejemplo 9.5 (f), estos conjuntos son cerrados en $\mathbf{R}^p$ y es claro que F_1 está acotado y que $F_1 \supseteq F_2 \supseteq \cdots \supseteq F_k \supseteq \cdots$.

Además, por la definición de d en F_k, se puede ver que F_k es no vacío. Del teorema de intersección de Cantor 11.4 se deduce que hay un punto y perteneciente a todos los F_k, $k \in \mathbf{N}$. Se puede ver fácilmente que $\|x-y\| = d$, de modo que y satisface la conclusión. (Fig. 11.3.) Q.E.D.

Una variante del siguiente teorema es de considerable importancia en la teoría de funciones analíticas. Se planteará el resultado únicamente para $p = 2$ y se emplearán ideas intuitivas al expresar lo que significa que un conjunto esté rodeado por una curva cerrada (es decir, una curva que no tiene puntos terminales).

11.7 TEOREMA DEL CONTORNO CIRCUNSCRIPTIVO. *Sea F un conjunto cerrado y acotado en $\mathbf{R}^2$ y sea G un conjunto abierto que contiene a F. Entonces, existe una curva cerrada C, que cae por completo en G y hecha de arcos de un número finito de círculos, tal que F está rodeado por C.*

DEMOSTRACION PARCIAL. Si x pertenece a $F \subseteq G$, existe un número $\delta(x) > 0$ tal que si $\|y-x\| < \delta(x)$, entonces y también pertenece a G. Sea $G(x) = \{y \in \mathbf{R}^2 : \|y-x\| < \frac{1}{2}\delta(x)\}$ para cada x en F. Dado que la colección $\mathcal{G} = \{G(x) : x \in F\}$ forma una cubierta del conjunto compacto F, la unión

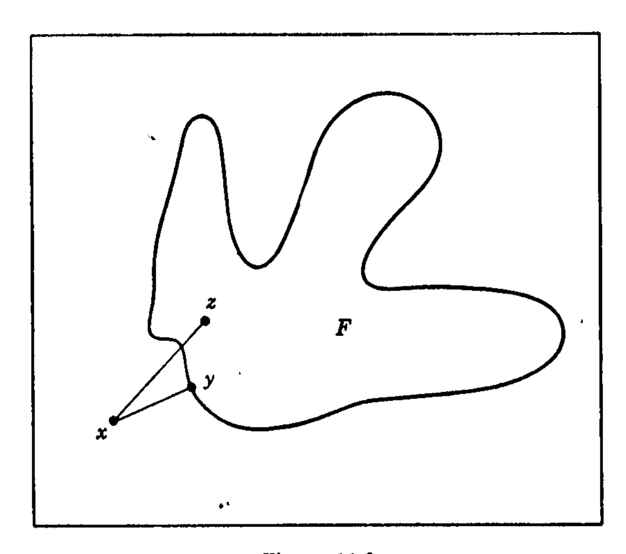


Figura 11.3

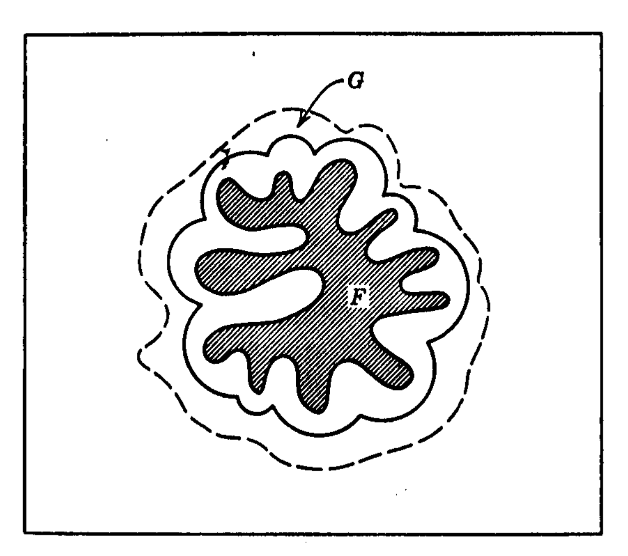


Figura 11.4

de un número finito de conjuntos en $\mathcal{G}$, digamos $G(x_1), \ldots, G(x_k)$, contiene al conjunto compacto F. Usando arcos de los círculos con centros x_j y radios $\frac{1}{2}\delta(x_j)$, se obtiene la curva deseada C. (Véase la Fig. 11.4.) La construcción detallada de la curva no se dará aquí.

Ejercicios

11.A. Demostrar, directamente de la definición (i.e., sin usar el teorema de Heine-Borel), que la bola abierta dada por $\{(x, y): x^2 + y^2 < 1\}$ no es compacta en R^2.

11.B. Demostrar directamente que la totalidad del espacio $\boldsymbol{R}^2$ no es compacta.

11.C. Demostrar directamente que si K es compacto en $\boldsymbol{R}^p$ y $F \subseteq K$ es un conjunto cerrado, entonces F es compacto en $\boldsymbol{R}^p$.

11.D. Demostrar que si K es un subconjunto compacto de $\boldsymbol{R}$, entonces K es compacto cuando se considera como un subconjunto de $\boldsymbol{R}^2$.

11.E. Modificando el argumento del ejemplo 11.2(d), demostrar que el intervalo $J = \{(x, y): 0 \leq x \leq 1, 0 \leq y \leq 1\}$ es compacto en $\boldsymbol{R}^2$.

11.F. Localizar los lugares en que se usaron las hipótesis de que el conjunto K es acotado y cerrado en la demostración del teorema de Heine-Borel.

11.G. Demostrar el teorema de intersección de Cantor eligiendo un punto x_n de F_n y después aplicando el teorema de Bolzano-Weierstrass 10.6 al conjunto $\{x_n : n \in \boldsymbol{N}\}$.

11.H. Si F es cerrado en $\boldsymbol{R}^p$ y si

$$d(x, F) = \inf \{\|x - z\| : z \in F\} = 0,$$

entonces x pertenece a F.

11.I. ¿El teorema del punto más próximo en $\boldsymbol{R}$ implica que hay un número real estrictamente positivo que es el más próximo a cero?

11.J. Si F es un conjunto no vacío cerrado en $\boldsymbol{R}^p$ y si $x \notin F$, ¿hay algún punto *único* de F que sea el más próximo a x?

11.K. Si K es un subconjunto compacto de $\boldsymbol{R}^p$ y x un punto de $\boldsymbol{R}^p$, entonces, el conjunto $K_x = \{x + y : y \in K\}$ también es compacto. (A este conjunto K_x algunas veces se le llama la **traslación** del conjunto K por medio de x.)

11.L. La intersección de dos conjuntos abiertos es compacto si y sólo si es vacía. ¿Puede ser la intersección de una colección infinita de conjuntos abiertos un conjunto compacto no vacío?

11.M. Si F es un subconjunto compacto de $\boldsymbol{R}^2$ y G es un conjunto abierto que contiene a F, entonces existe una curva poligonal cerrada C que está enteramente en G que rodea a F.

11.N. Sea $\{H_n : n \in \boldsymbol{N}\}$ una familia de subconjuntos cerrados de $\boldsymbol{R}^p$ con la propiedad de que ningún conjunto H_n contiene algún conjunto abierto no vacío (por ejemplo, H_n es un punto o una línea en $\boldsymbol{R}^2$). Sea $G \neq \emptyset$ un conjunto abierto.

(a) Si $x_1 \in G \setminus H_1$, demostrar que existe una bola cerrada B_1 con centro x_1 tal que $B_1 \subseteq G$ y $H_1 \cap B_1 = \emptyset$.

(b) Si $x_2 \notin H_2$ pertenece al interior de B_1, demostrar que existe una bola cerrada B_2 con centro x_2 tal que B_2 está contenido en el interior de B_1 y $H_2 \cap B_2 = \emptyset$.

(c) Continuar este proceso para obtener una familia nidificada de bolas cerradas tal que $H_n \cap B_n = \emptyset$. Por el teorema de intersección de Cantor 11.4 se sabe que hay un punto x_0 común a todas las B_n. Deducir que $x_0 \in G \setminus \bigcap H_n$, de manera que G no puede estar contenido en $\bigcup H_n$. Este resultado es una forma de lo que a menudo se llama "el teorema de categoría de Baire†".

11.O. Una **línea** en $\boldsymbol{R}^2$ es un conjunto de puntos (x, y) que satisfacen una ecuación, de la forma $ax + by + c = 0$ en donde $(a, b) \neq (0, 0)$. Emplear el ejercicio anterior para demostrar que $\boldsymbol{R}^2$ no es la unión de una colección contable de líneas.

11.P. El conjunto $\mathscr{C}(\boldsymbol{Q})$ de números irracionales en $\boldsymbol{R}$ no es la unión de una familia contable de conjuntos cerrados, ninguno de los cuales contiene algún conjunto abierto no vacío.

11.Q. El conjunto $\boldsymbol{Q}$ de números racionales no es la intersección de una colección contable de conjuntos abiertos en $\boldsymbol{R}$.

Sección 12 Conjuntos conexos

En seguida se introducirá el concepto de conjunto conexo que se empleará en ocasiones en lo sucesivo.

12.1. DEFINICION. Se dice que un subconjunto $D \subseteq \boldsymbol{R}^p$ es **inconexo** si existen dos conjuntos abiertos A, B tales que $A \cap D$ y $B \cap D$ son ajenos, no vacíos y su unión es D. En este caso se dice que el par A, B forma una **inconexión** de D. Un subconjunto que no es inconexo se dice que es **conexo.** (Fig. 12.1.)

12.2 EJEMPLOS. *(a)* El conjunto $\boldsymbol{N} \subseteq \boldsymbol{R}$ es inconexo ya que se pueden tomar $A = \{x \in \boldsymbol{R} : x < 3/2\}$ y $B = \{x \in \boldsymbol{R} : x > 3/2\}$.

(b) El conjunto $H = \{1/n : n \in \boldsymbol{N}\}$ es inconexo.

†RENE LOUIS BAIRE (1874-1932) fue profesor en Dijon. Trabajó en teoría de conjuntos y en análisis real.

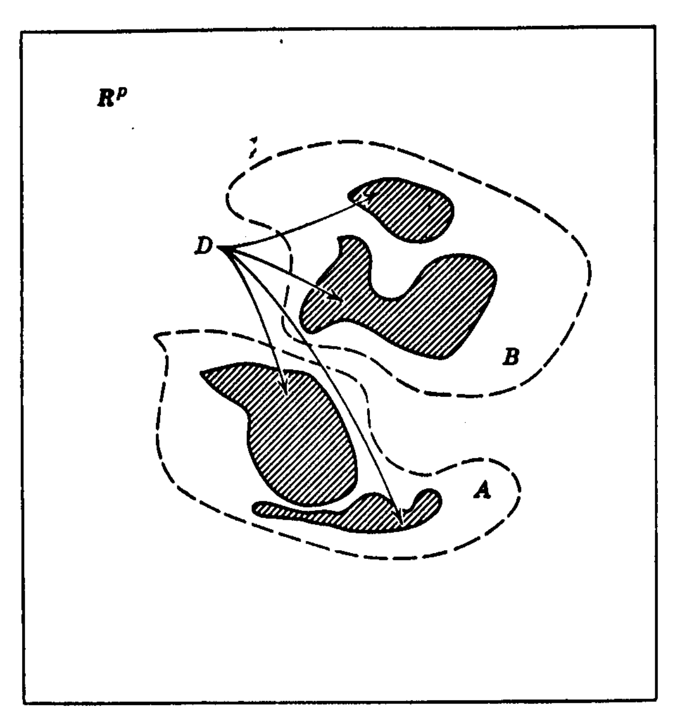


Figura 12.1. Un conjunto inconexo.

(c) El conjunto S que consta de todos los números racionales positivos es inconexo en $\boldsymbol{R}$ ya que se pueden tomar $A=\{x\in\boldsymbol{R}: x<\sqrt{2}\}$ y $B=\{x\in\boldsymbol{R}: x>\sqrt{2}\}$.

(d) Si $0<c<1$, entonces los conjuntos $A=\{x\in\boldsymbol{R}, x\leq c\}$, $B=\{x\in\boldsymbol{R}: x>c\}$ dividen al intervalo unitario $\boldsymbol{I}=\{x\in\boldsymbol{R}: 0\leq x\leq 1\}$ en conjuntos ajenos, no vacíos y cuya unión es $\boldsymbol{I}$. Sin embargo, dado que A no es abierto, este ejemplo *no* prueba que $\boldsymbol{I}$ sea inconexo. En seguida se demostrará que, de hecho, el conjunto $\boldsymbol{I}$ es conexo.

12.3 TEOREMA. *El intervalo unitario cerrado* $\boldsymbol{I}=[0, 1]$ *es un subconjunto conexo de* $\boldsymbol{R}$.

DEMOSTRACION. Se llevará a cabo por contradicción suponiendo que A y B son conjuntos abiertos que forman una inconexión de $\boldsymbol{I}$. De modo que $A\cap\boldsymbol{I}$ y $B\cap\boldsymbol{I}$ son conjuntos ajenos no vacíos acotados cuya unión es $\boldsymbol{I}$. Dado que los conjuntos A y B son abiertos, los conjuntos $A\cap\boldsymbol{I}$ y $B\cap\boldsymbol{I}$ no pueden constar sólo de un pun to. (¿por qué?) Para ser más precisos, suponga que existen puntos $a\in A$, $b\in B$ tales que $0<a<b<1$. Aplicando la propiedad del supremo 6.4, suponga $c=\sup\{x\in A: x<b\}$ de manera que $0<c<1$; por lo tanto, $c\in A\cup B$. Si $c\in A$, ntonces $c\neq b$ y como A es abierto hay un punto $a_1\in A$, $c<a_1$, tal que el intervalo $[c, a_1]$ está contenido en $\{x\in A: x<b\}$, contrario a la definición de c. Análogamente, si $c\in B$, entonces, puesto que B es abierto, hay un punto $b_1\in B$, $b_1<c$, tal que el intervalo

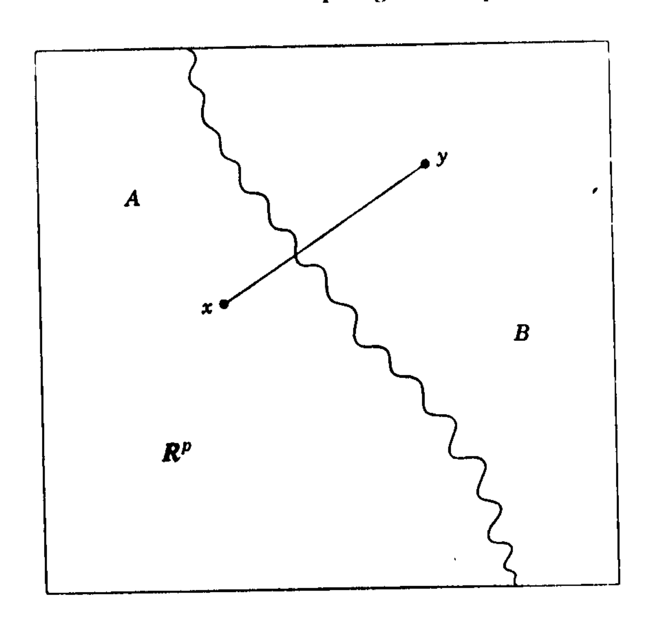


Figura 12.2

$[b_1, c]$ está contenido en $B \cap I$, contrario a la definición de c. Por lo tanto, la hipótesis de que I es inconexo induce una contradicción. Q.E.D.

El lector deberá notar que la misma demostración se puede usar para probar que el intervalo abierto (0,1) es conexo en R.

12.4 TEOREMA. *La totalidad del espacio* R^p *es conexo.*

DEMOSTRACION. De no ser así, existirían dos conjuntos abiertos ajenos no vacíos A, B cuya unión sería R^p. (Fig. 12.2) Sean $x \in A$ y $y \in B$ y considere al segmento de línea S que une a x con y; es decir,

$$S = \{x + t(y - x) : t \in I\}.$$

Sean $A_1 = \{t \in R : x + t(y - x) \in A\}$ y $B_1 = \{t \in R : x + t(y - x) \in B\}$; fácilmente se puede ver que A_1 y B_1 son subconjuntos abiertos ajenos no vacíos de R y proporcionan una inconexión de I, contradiciéndose el teorema 12.3
Q.E.D.

12.5 COROLARIO. *Los únicos subconjuntos de* R^p *que son tanto abiertos como cerrados son* $\emptyset$ *y* R^p.

DEMOSTRACION. Si R^p, es tanto abierto como cerrado en R^p, entonces $B = R^p \setminus A$ también lo es. Si A no es vacío y no es todo R^p, entonces el par A, B forma una inconexión de R^p, contradiciéndose el teorema.
Q.E.D.

Conjuntos abiertos conexos

En ciertas áreas del análisis los conjuntos abiertos conexos desempeñan un papel especialmente importante. Usando la definición es fácil probar el siguiente resultado.

12.6 LEMA. *Un subconjunto abierto de* $\mathbf{R}^p$ *es conexo si y sólo si no se puede expresar como la unión de dos conjuntos abiertos ajenos no vacíos.*

En algunas ocasiones es útil tener otra caracterización de conjuntos abiertos conexos. Para poder dar dicha caracterización se debe introducir cierta terminología. Si x y y son dos puntos en $\mathbf{R}^p$, entonces una **curva poligonal** que une a x con y es un conjunto P que se obtiene de la unión de un número finito de segmentos de línea ordenados $(L_1, L_2, \ldots, L_n)$ en $\mathbf{R}^p$, tales que el segmento de línea L_1 tiene como puntos terminales x, z_1; el segmento de línea L_2 tiene como puntos terminales z_1, z_2; ... ; y el segmento de línea L_n tiene como puntos terminales z_{n-1}, y. (Fig. 12.3.)

12.7 TEOREMA. *Sea G un conjunto abierto en* $\mathbf{R}^p$. *G es conexo si y sólo si cualquier par de puntos x,y en G se puede unir por medio de una curva poligonal que cae enteramente en G.*

DEMOSTRACION. Suponga que G no es conexo y que A,B es una inconexión de G. Sean $x \in A \cap G$ y $y \in B \cap G$ y sea $P = (L_1, L_2, \ldots, L_n)$ una curva poligonal que cae enteramente en G que une a x con y. Sea k el número natural más pequeño tal que el punto terminal z_{k-1} de L_k pertenece a $A \cap G$ el punto terminal z_k pertenece a $B \cap G$ (Fig. 12.4). Si A_1 y B_1 se definen como

$$A_1 = \{t \in \mathbf{R} : z_{k-1} + t(z_k - z_{k-1}) \in A \cap G\},$$
$$B_1 = \{t \in \mathbf{R} : z_{k-1} + t(z_k - z_{k-1}) \in B \cap G\},$$

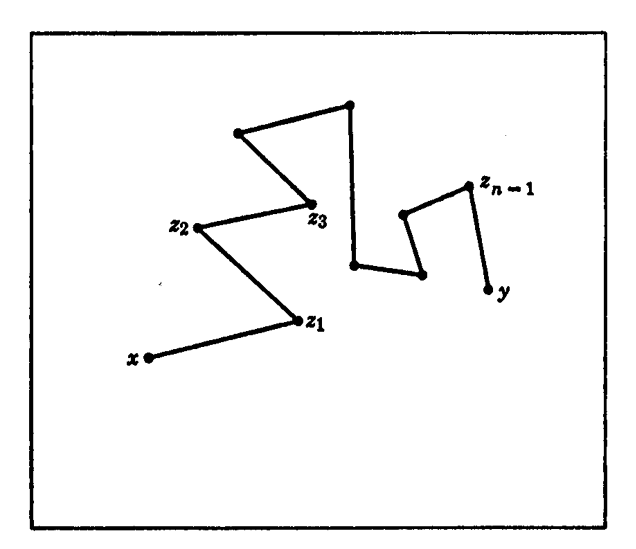


Figura 12.3. Una curva poligonal.

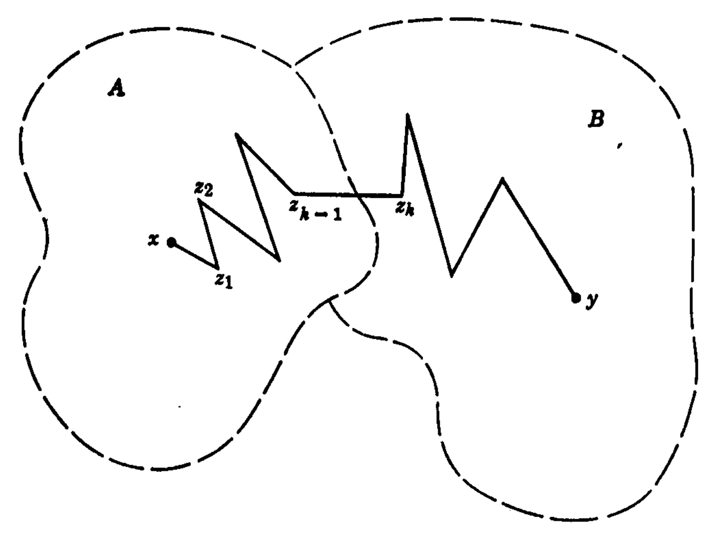


Figura 12.4.

entonces se puede ver fácilmente que A_1 y B_1 son subconjuntos abiertos ajenos no vacíos de **R**. De donde, el par A_1, B_1 forma una inconexión del intervalo unitario **I**, contradiciéndose el teorema 12.3. Por lo tanto, si G no es conexo, existen dos puntos en G que no se pueden unir por medio de una curva poligonal en G.

Ahora, suponga que G es un conjunto abierto conexo en $\boldsymbol{R}^p$ y que x pertenece a G. Sea G_1 el subconjunto de G que consta de todos los puntos de G que se pueden unir a x por medio de una curva poligonal que cae enteramente en G. Sea G_2 tal que conste de todos los puntos de G que no se pueden unir a x por medio de una curva poligonal que cae en G. Es claro que $G_1 \cap G_2 = \emptyset$. El conjunto G_1 no es vacío ya que contiene al punto x. Ahora se demostrará que G_1 es abierto en $\boldsymbol{R}^p$. Si y pertenece a G_1, del hecho de que G es abierto se deduce que para algún número real $r > 0$, $\|w - y\| < r$ implica que $w \in G$. Por la definición de G_1, el punto y se puede unir a x por medio de una curva poligonal, y añadiendo un segmento de y a w se deduce que w pertenece a G_1. Por lo tanto, G_1 es un subconjunto abierto de $\boldsymbol{R}^p$. Análogamente, el subconjunto G_2 es abierto en $\boldsymbol{R}^p$. Si G_2 no es vacío, los conjuntos G_1, G_2 forman una inconexión de G, contrario a la hipótesis de que G es conexo. Por lo tanto, $G_2 = \emptyset$ y todo punto de G se puede unir a x por medio de una curva poligonal que cae por completo en G.

Conjuntos conexos en *R*

Se concluye esta sección demostrando que los subconjuntos conexos de $\boldsymbol{R}$ son precisamente los intervalos (véase la sección 7).

12.8 TEOREMA. *Un subconjunto* $\boldsymbol{R}$ *es conexo si y sólo si es un intervalo.*

DEMOSTRACION PARCIAL. La demostración que se da en el teorema 12.3 se puede modificar fácilmente para probar la conexidad de un intervalo arbitrario no vacío. Se le dejan al lector los detalles.

A la inversa, sea $C \subseteq \boldsymbol{R}$ conexo y suponga que $C \neq \emptyset$. Observe que C tiene la propiedad de que si a, $b \in C$ y $a < b$, entonces cualquier número c que satisfaga $a < c < b$ también debe pertenecer a C, ya que si $c \notin C$, los conjuntos $A = \{x \in \boldsymbol{R} : x < c\}$ y $B = \{x \in \boldsymbol{R} : x > c\}$ forman una inconexión de C.

(i) Suponga ahora que C está acotado por arriba y por abajo y sean $a = \inf C$ y $b = \sup C$. Se probará que C debe tener una de las cuatro formas

$$[a, b], \qquad [a, b), \qquad (a, b], \qquad (a, b).$$

De hecho, si $a \in C$ y $b \in C$, se ha visto en el párrafo anterior que $[a, b] \subseteq C$ y el hecho de que $C \subseteq [a, b]$ se deduce del hecho de que a y b son cotas inferior y superior, respectivamente, de C.

Si $a \in C$ pero $b \notin C$, sea b' cualquier número con $a \leq b' < b$. Dado que $b = \sup C$, debe haber una elemento $b'' \in C$ tal que $a \leq b' < b''$. Por lo tanto, el número b' debe pertenecer a C y ya que b' es cualquier número que satisface $a \leq b' < b$, se deduce que $C = [a, b)$.

Análogamente, si $a \notin C$ pero $b \in C$, se deduce que $C = (a, b]$, mientras que si $a \notin C$ y $b \notin C$, entonces se deduce que $C = (a, b)$.

(ii) Suponga ahora que C está acotado por abajo pero no por arriba y sea $a = \inf C$, de manera que $C \subseteq [a, +\infty)$. Si $a \in C$ y si x es cualquier núm ero real con $a \leq x$, entonces como C no está acotado por arriba, existe $c \in C$ tal que $x \leq c$ por lo que de la propiedad anterior se deduce que $x \in C$. Como x es un número arbitrario que satisface $a \leq x$, se concluye que $C = [a, +\infty)$.

Análogamente, si $a \notin C$ se concluye que $C = (a, +\infty)$.

(iii) Si C no está acotado por abajo pero está acotado por arriba y si $b = \sup C$, entonces hay dos casos: $C = (-\infty, b]$ ó $C = (-\infty, b)$ según sea $b \in C$ ó $b \notin C$.

(iv) Por último, si C no está acotado ni por abajo ni por arriba, entonces se tiene el caso $C = (-\infty, +\infty)$.

Ejercicios

12.A. Si A y B son subconjuntos conexos de $\boldsymbol{R}^p$, dar ejemplos para demostrar que $A \cup B$, $A \cap B$, $A \setminus B$ pueden ser conexos o inconexos.

12.B. Si $C \subseteq \boldsymbol{R}^p$ es conexo y x es un punto de acumulación de C, entonces $C \cup \{x\}$ es conexo.

12.C. Si $C \subseteq \boldsymbol{R}^p$ es conexo, demostrar que su cerradura C^- (véase el ejercicio 9.L) también es conexa.

12.E. Si $K \subseteq \boldsymbol{R}^p$ es convexo (véase el ejercicio 8.Q), entonces, K es conexo.

12.F. El conjunto de Cantor $\boldsymbol{F}$ es inconexo inordenadamente. Demostrar que si x, $y \in \boldsymbol{F}$, $x \neq y$, entonces hay una inconexión A, B de $\boldsymbol{F}$ tal que $x \in A$, $y \in B$.

12.G. Si C_1 y C_2 son subconjuntos conexos de $\mathbf{R}$, entonces el producto $C_1 \times C_2$ es un subconjunto conexo de $\boldsymbol{R}^2$.

12.H. Demostrar que el conjunto

$$A = \{(x, y) \in \boldsymbol{R}^2 : 0 < y \leq x^2, x \neq 0\} \cup \{(0, 0)\}$$

es conexo en $\boldsymbol{R}^2$. Sin embargo, no existe ninguna curva poligonal que cae por completo en A que una a (0,0) con otros puntos del conjunto.

12.I Demostrar que el conjunto

$$S = \left\{(x, y) \in \boldsymbol{R}^2 : y = \operatorname{sen} \frac{1}{x}, x \neq 0\right\} \cup \left\{(0, y) : -1 \leq y \leq 1\right\}$$

es conexo en $\boldsymbol{R}^2$. Sin embargo, no siempre es posible unir dos puntos en S por medio de una curva poligonal (o cualquier curva "continua") que cae por completo en S.

Sección 13 El sistema de números complejos

Teniendo el sistema de números reales a la mano, resulta fácil crear el sistema de números complejos. En esta sección se indicará cómo se puede construir el campo de los complejos†.

Como se vio antes, el sistema de números reales es un campo que satisface ciertas propiedades adicionales. En la sección 8 se construyó el espacio cartesiano $\boldsymbol{R}^p$ y se introdujeron algunas operaciones en el producto cartesiano de p copias de $\boldsymbol{R}$. sin embargo, $\boldsymbol{R}^p$ no se hizo un campo. Puede resultar extraño el hecho de que no sea posible definir una multiplicación que convierta a $\boldsymbol{R}^p, p \geq 3$, en un campo. No obstante, es posible definir una operación de multiplicación en $\boldsymbol{R} \times \boldsymbol{R}$ que haga de este conjunto un campo. En seguida se introducirán las operaciones deseadas.

13.1 DEFINICION. El sistema de números complejos $\boldsymbol{C}$ consta de todos los pares ordenados (x,y) de números reales con la operación de **adición** definida por

$$(x, y) + (x', y') = (x + x', y + y'),$$

y la operación de **multiplicación** definida por

$$(x, y) \cdot (x', y') = (xx' - yy', xy' + x'y).$$

De manera que el sistema de números complejos $\boldsymbol{C}$ tiene los mismos elementos que el espacio bidimensional $\boldsymbol{R}^2$. Tiene la misma operación de adición pero posee una multiplicación que $\boldsymbol{R}^2$ no tiene. Por lo tanto, considerados

†Esta sección se puede omitir en una primera lectura.

simplemente como conjuntos, C y R^2 son iguales ya que tienene los mismos elementos; sin embargo, desde el punto de vista algebraico no son lo mismo ya que poseen operaciones distintas.

A un elemento de C se le llama un **número complejo** y a menudo se denota mediante una sola letra como z. Si $z = (x, y)$, se dice que el número real x es la **parte real** de z y que y es la **parte imaginaria** de z, por medio de símbolos,

$$x = \text{Re } z, \qquad y = \text{Im } z.$$

Al número complejo $\bar{z} = (x, -y)$ se le llama el **conjugado** de $z = (x, y)$.

Resulta un hecho importante que las definiciones que se acaban de dar acerca de suma y multiplicación de elementos de C lo hagan un "campo" en el sentido del álgebra abstracta. Es decir, que satisface las propiedades algebraicas enunciadas en 4.1, tomando en cuenta que el número 0 en (A3) se reemplaza por el par (0,0), el elemento correspondiente a $-a$ en (A4) es el par $(-x, -y)$, el número 1 en (M3) se reemplaza por el par (1,0) y el número correspondiente a $1/a$ es el par

$$\left(\frac{x}{x^2+y^2}, \quad \frac{-y}{x^2+y^2}\right)$$

cuando $(x, y) \neq (0, 0)$.

Algunas veces es conveniente adoptar parte de la notación de la sección 8 y escribir

$$az = a(x, y) = (ax, ay),$$

cuando a es un número real y $z = (x, y)$ está en C. Es claro que con esta notación cada elemento en C tiene una representación única como la suma del producto de un número real por (1,0) y el producto de un número real por (0,1). De modo que se puede escribir

$$z = (x, y) = x(1, 0) + y(0, 1).$$

Dado que el elemento (1,0) es el elemento identidad de C, es natural que se designe con 1 (o que se suprima por completo cuando aparece como factor). Para abreviar un poco, es conveniente introducir un símbolo para (0,1) y la elección convenida es i. Usando esta notación se escribe

$$z = (x, y) = x + iy.$$

Además, se tiene $\bar{z} = (x, -y) = x - iy$ y

$$x = \text{Re } z = \frac{z+\bar{z}}{2}, \qquad y = \text{Im } z = \frac{z-\bar{z}}{2i}$$

Por la definición 13.1 se tiene $(0, 1)(0, 1) = (-1, 0)$ que se puede escribir como $i^2 = -1$. De modo que en C la ecuación cuadrática

$$z^2 + 1 = 0,$$

tiene solución. El motivo histórico por el que se desarrolló el sistema de números complejos fue la obtención de un sistema de "números" en el que toda ecuación cuadrática tuviera solución. Se observó que no toda ecuación con coeficientes reales tiene una solución real y para remediar esta deficiencia se inventaron los números complejos. Es bien sabido que los números complejos no sólo producen soluciones para ecuaciones cuadráticas con coeficientes reales sino que también garantizan soluciones de ecuaciones polinomiales de cualquier grado y con coeficientes que pueden ser números complejos. A este resultado se le llama el teorema fundamental del álgebra y lo demostró por primera vez Gauss† en 1799.

A pesar de que a C no se le pueden adjudicar las propiedades analizadas en la sección 5, es fácil dotarlo de la estructura métrica y topológica de las secciones 8 y 9. Ya que si $z = (x, y)$ pertenece a C, el *valor absoluto* de z se define como

$$|z| = (x^2 + y^2)^{1/2}.$$

Es fácil ver que el valor absoluto que se acaba de definir tiene las siguientes propiedades:

(i) $|z| \geq 0$; si y sólo si
(ii) $|z| = 0$
(iii) $|wz| = |w|\,|z|$;
(iv) $\big||w| - |z|\big| \leq |w \pm z| \leq |w| + |z|$.

Se observará que el valor absoluto del número complejo $z = (x, y)$ es precisamente el mismo de la norma de un elemento (x, y) en $\boldsymbol{R}^2$. Por lo tanto, todas las propiedades topológicas de los espacios cartesianos que se introdujeron y estudiaron de la sección 9 a la 12 tiene sentido y son válidas para C. En particular, los conceptos de conjuntos abiertos y cerrados en C son exactamente iguales que para el espacio cartesiano $\boldsymbol{R}^2$. Más aún, el teorema de Bolzano-Weierstrass 10.6 y el teorema de Heine-Borel 11.3, junto con sus consecuencias, también son válidos en C, lo mismo que el teorema 12.7.

El lector deberá tener presentes estas observaciones en todas las secciones restantes del texto. Se verá que *todo el material que sigue, aplicable a espacios cartesianos de dimensión mayor a uno, se aplica igualmente al sistema de números complejos.* De manera que la mayoría de los resultados que se obtengan acerca de sucesiones, funciones continuas, derivadas, integrales y series infinitas también son válidas para C, sin cambiar el enunciado ni la de-

†CARL FRIEDRICH GAUSS (1777-1855), el prodigioso hijo de un jornalero, fue uno de los más grandes matemáticos; sin embargo, también se le recuerda por sus trabajos en astronomía, física y geodesia. Fue profesor y director del Observatorio en Gottingen.

mostración. Las únicas excepciones a esta afirmación son las propiedades basadas en las propiedades de orden de $\boldsymbol{R}$.

En este sentido, el análisis complejo es un caso especial del análisis real; sin embargo, hay varias propiedades nuevas que son profundas e importantes en el estudio de funciones analíticas que no tienen ningún equivalente en el ámbito del análisis real. Por lo tanto, sólo los aspectos bastante superficiales del análisis complejo se incluyen en lo que se habrá de hacer.

Ejercicios

13.A. Demostrar que el número complejo iz se obtiene de z mediante una rotación, en el sentido opuesto al de las manecillas del reloj de $\pi/2$ radianes (= 90°) en torno al origen.

13.B. Si $c = (\cos\theta, \operatorname{sen}\theta) = \cos\theta + i \operatorname{sen}\theta$, entonces el número cz se obtiene de z mediante una rotación en el sentido opuesto al de las manecillas del reloj, de θ radianes en torno al origen.

13.C. Describir la relación geométrica entre los números complejos z y $az+b$, en donde $a \neq 0$. Demostrar que la aplicación definida para $z \in \boldsymbol{C}$, por medio de $f(z) = az + b$, manda círculos a círculos y líneas a líneas.

13.D. Describir las relaciones geométricas entre los números complejos z, $\bar{z}$ y $1/z$ para $z \neq 0$. Demostrar que la aplicación definida por medio de $g(z) = \bar{z}$ manda círculos a círculos y líneas a líneas. ¿Cuales círculos y líneas permanecen fijos bajo g?

13.E. Demostrar que la **aplicación inversión** definida por $h(z) = 1/z$, manda círculos y líneas a círculos y líneas. ¿Cuáles círculos se mandan a líneas? ¿cuáles líneas se mandan a círculos? Examinar las imágenes bajo h de las líneas verticales dadas por la ecuación $\operatorname{Re} z$ = constante, las líneas horizontales $\operatorname{Im} z$ = constante y los círculos $|z|$ = constante.

13.F. Investigar el carácter geométrico de la aplicación definida por $g(z) = z^2$. Determinar si la aplicación g es uno a uno y si aplica a $\boldsymbol{C}$ sobre todo $\boldsymbol{C}$. Examinar las imágenes inversas bajo g de las líneas

$$\operatorname{Re} z = \text{constante}, \qquad \operatorname{Im} z = \text{constante},$$

y los círculos $|z|$ = constante.

III CONVERGENCIA

El material de los dos capítulos anteriores debe proporcionar una comprensión adecuada del sistema de números reales y de espacios cartesianos. Una vez expuestos estos fundamentos algebraicos y topológicos se puede proceder a plantear preguntas de naturaleza más analítica. Se empezará por un estudio acerca de la convergencia de sucesiones. Algunos de los resultados de este capítulo posiblemente le sean familiares al lector debido a otros cursos de análisis, pero se pretende que la presentación que aquí se da sea más rigurosa y que se obtengan resultados más a fondo que en cursos anteriores.

Primero se dará el significado de convergencia de una sucesión de elementos en $\boldsymbol{R}^p$ y se establecerán algunos resultados elementales (pero útiles) acerca de sucesiones convergentes. Después se establecerán algunos criterios importantes en cuanto a convergencia. En seguida se estudia la convergencia y la convergencia uniforme de sucesiones de funciones. Luego de una breve sección en relación con el límite superior, se anexa una última sección que, a pesar de ser interesante, se puede omitir sin pérdida de continuidad, ya que esos resultados no se aplicarán más adelante.

Debido a las limitaciones que siempre existen en un libro, se ha decidido seguir este capítulo con un estudio de continuidad, diferenciación e integración. Desafortunadamente esto aplaza la presentación completa de series más adelante. Se recomienda al instructor dar, cuando menos, una breve introducción de series en forma simultánea con este capítulo o si lo prefiere puede pasar directamente a la primera parte del capítulo IV al terminar la sección 16.

Sección 14 Introducción a las sucesiones

A pesar de que la teoría de convergencia se puede dar a un nivel muy abstracto es preferible estudiar la convergencia de sucesiones en un espacio cartesiano $\boldsymbol{R}^p$, prestando especial atención al caso de la recta real. El lector deberá interpretar las ideas dibujando diagramas en $\boldsymbol{R}$ y $\boldsymbol{R}^2$

14.1 DEFINICION. Si S es cualquier conjunto, una **sucesión** en S es una función sobre el conjunto $\mathbf{N}=\{1, 2, \ldots\}$ de números naturales y cuyo rango está en S. En particular una sucesión en $\boldsymbol{R}^p$ es una función cuyo dominio es N y cuyo rango está contenido en $\boldsymbol{R}^p$.

En otras palabras una sucesión en $\boldsymbol{R}^p$ asigna a cada número natural $n = 1, 2, \ldots,$ un elemento de $\boldsymbol{R}^p$ determinado de manera única. Tradicionalmente el elemento $\boldsymbol{R}^p$ que se asigna a un número natural n se denota mediante algún símbolo como x_n a pesar de que esta notación varía con respecto a las que se emplean para la mayoría de las funciones, se habrá de utilizar la simbología convencional. [De acuerdo con la notación anterior si $X: \mathbf{N} \to \boldsymbol{R}^p$ es una sucesión el valor de X en $n \in \mathbf{N}$ se debe representar por $X(n)$ y no por x_n.]

Si bien se acepta la notación tradicional, también se quiere hacer la distinción entre la función X y sus valores $X(n) = x_n$. De modo que cuando los elementos de la sucesión (es decir los valores de la función) se designan con x_n, la función se designará por medio de $X = (x_n)$ o $X = (x_n : n \in \mathbf{N})$. Se usan paréntesis para indicar que el orden inducido por el de N es una cuestión importante. De modo que por medio de la notación se está distinguiendo entre la sucesión $X = (x_n : n \in \mathbf{N})$ y el conjunto $\{x_n : n \in \mathbf{N}\}$ de valores de esta sucesión.

Al definir sucesiones a menudo se hace una lista ordenada de los elementos de la sucesión deteniéndose cuando la regla de construcción resulte evidente. De modo que se puede escribir

$$(2, 4, 6, 8, \ldots)$$

para representar a la sucesión de enteros pares. Un método más eficaz resulta al dar una fórmula específica para el término general de la sucesión tal como

$$(2n : n \in \mathbf{N}).$$

En la práctica con frecuencia es más conveniente especificar el valor x_1 y algún método para obtener x_{n+1}, $n \geq 1$, cuando x_n se conoce. Generalizando aún más se pueden hacer explícitos x_1 y alguna regla para obtener x_{n+1} a partir de $x_1, x_2, \ldots, x_n$. Se hará referencia a cualquiera de estos métodos como las definiciones *inductivas* de la sucesión. De esta manera se podría definir la sucesión de números naturales pares por medio de la definición

$$x_1 = 2, \qquad x_{n+1} = x_n + 2, \qquad n \geq 1.$$

o por la definición (al parecer más complicada)

$$x_1 = 2, \qquad x_{n+1} = x_n + x_1, \qquad n \geq 1.$$

Es claro que para definir esta sucesión existe la posibilidad de usar muchos otros métodos.

En seguida se introducen algunos métodos para construir sucesiones nuevas a partir de otras dadas.

14.2 DEFINICION. Si $X = (x_n)$ y $Y = (y_n)$ son sucesiones en $\boldsymbol{R}^p$, la **suma** de estas sucesiones se define como la sucesión $X + Y = (x_n + y_n)$ en $\boldsymbol{R}^p$,

su **diferencia** como la sucesión $X - Y = (x_n - y_n)$ y su **producto interno** como la sucesión $X \cdot Y = (x_n \cdot y_n)$ en $\boldsymbol{R}$ que se obtiene mediante el producto interno de términos correspondientes. Análogamente, si $X = (x_n)$ es una sucesion en $\boldsymbol{R}$ y si $Y = (y_n)$ es una sucesión en $\boldsymbol{R}^p$, el **producto** de X y Y se define como la sucesión en $\boldsymbol{R}^p$ denotada mediante $XY = (x_n y_n)$; o si $c \in \boldsymbol{R}$ y $X = (x_n)$, se define $cX = (cx_n)$. Por último, si $Y = (y_n)$ es una sucesión en $\boldsymbol{R}$ con $y_n \neq 0$, se puede definir al **cociente** de la sucesión $X = (x_n)$ en $\boldsymbol{R}^p$ entre Y como la sucesión $X/Y = (x_n/y_n)$.

Por ejemplo si X, Y son las sucesiones en $\boldsymbol{R}$ dadas por

$$X = (2, 4, 6, \ldots, 2n, \ldots), \qquad Y = \left(1, \frac{1}{2}, \frac{1}{3}, \ldots, \frac{1}{n}, \ldots\right),$$

entonces se tiene

$$X + Y = \left(3, \frac{9}{2}, \frac{19}{3}, \ldots, \frac{2n^2+1}{n}, \ldots\right),$$

$$X - Y = \left(1, \frac{7}{2}, \frac{17}{3}, \ldots, \frac{2n^2-1}{n}, \ldots\right),$$

$$XY = (2, 2, 2, \ldots, 2, \ldots),$$

$$3X = (6, 12, 18, \ldots, 6n, \ldots),$$

$$\frac{X}{Y} = (2, 8, 18, \ldots, 2n^2, \ldots).$$

Análogamente si Z denota la sucesión en H dada por

$$Z = \left(1, 0, 1, \ldots, \frac{1-(-1)^n}{2}, \ldots\right),$$

ya se han definido $X + Z$, $X - Z$ y XZ; sin embargo X/Z no está definida ya que algunos de los elementos en Z son cero.

En seguida se dará el concepto del límite de una sucesión.

14.3 DEFINICION. Sea $X = (x_n)$ una sucesión en $\boldsymbol{R}^p$. Se dice que un elemento x de $\boldsymbol{R}^p$ es un **límite** de X si para cada vecindad V de x hay un número natural K_V tal que para toda $n \geq K_V$, x_n pertenece a V. Si x es un límite de X también se dice que X **converge** a x. Si una sucesión tiene un límite se dice que la sucesión es **convergente.** Si una sucesión no tiene ningún límite se dice que es **divergente.**

La notación K_V se usa para indicar que la elección de K dependerá de V. Es claro que una vecindad pequeña V por lo general requerirá un valor grande para poder garantizar que $x_n \in V$ para toda $n \geq K_V$.

Se ha definido el límite de una sucesión $X = (x_n)$ en términos de vecindades. A menudo es conveniente usar la norma en $\boldsymbol{R}^p$ para dar una definición equivalente la cual se dará en seguida mediante un teorema.

14.4 TEOREMA. *Sea $X = (x_n)$ una sucesión en $\mathbf{R}^p$. Un elemento x de $\mathbf{R}^p$ es un límite de X si y sólo si para cada $\varepsilon > 0$ hay un número natural $K(\varepsilon)$ tal que para toda $n \geq K(\varepsilon)$, $\|x_n - x\| < \varepsilon$.*

DEMOSTRACION. Suponga que x es un límite de la sucesión X, según la definición 14.3. Sea $\varepsilon > 0$ y considere a la bola abierta $V(\varepsilon) = \{y \in \mathbf{R}^p : \|y - x\| < \varepsilon\}$, que es una vecindad de x. Por la definición 14.3 se sabe que hay un número natural $K_{V(\varepsilon)}$ tal que si $n \geq K_{V(\varepsilon)}$, entonces $x_n \in V(\varepsilon)$. De donde si $n \geq K_{V(\varepsilon)}$, entonces $\|x_n - x\| < \varepsilon$. Esto prueba que la propiedad establecida es válida cuando x es un límite de X.

A la inversa suponga que la propiedad del teorema es válida para toda $\varepsilon > 0$; se debe probar que la definición 14.3 se satisface. Para hacerlo sea V cualquier vecindad de x; entonces hay un número $\varepsilon > 0$ tal que la bola abierta $V(\varepsilon)$ con centro x y radio x está contenida en V. De acuerdo con la propiedad del teorema, hay un número natural $K(\varepsilon)$ tal que si $n \geq K(\varepsilon)$, entonces $\|x_n - x\| < \varepsilon$. Expresado de otra manera si $n \geq K(\varepsilon)$, entonces $x_n \in V(\varepsilon)$; por lo tanto $x_n \in V$ y el requisito de la definición 14.3 se satisface. Q.E.D.

14.5 UNICIDAD DE LIMITES. *Una sucesión en $\mathbf{R}^p$ puede tener cuando más un límite.*

DEMOSTRACION. Suponga por el contrario, que x', x'' son límites de $X = (x_n)$ y que $x' \neq x''$. Sean V' y V'' vecindades ajenas de x' y x'' respectivamente y sean K', K'' números naturales tales que si $n \geq K'$ entonces $x_n \in V'$ y si $n \geq K''$ entonces $x_n \in V''$. Sea $K = \sup\{K', K''\}$ de manera que $x_K \in V'$ y $x_K \in V''$. Se deduce que x_K pertenece a $V' \cap V''$, contrario al supuesto de que V' y V'' son ajenos. Q.E.D.

Cuando una sucesión $X = (x_n)$ en $\mathbf{R}^p$ tiene un límite x a menudo se escribe

$$x = \lim X, \qquad \text{ó} \qquad x = \lim_n (x_n),$$

o se usa el simbolismo $x_n \rightarrow x$.

Se dice que una sucesión $X = (x_n)$ en $\mathbf{R}^p$ es **acotada** si existe $M > 0$ tal que $\|x_n\| < M$ para toda $n \in \mathbf{N}$.

14.6 LEMA. *Una sucesión convergente en $\mathbf{R}^p$ es acotada.*

DEMOSTRACION. Sea $x = \lim (x_n)$ y sea $\varepsilon = 1$. Por el teorema 14.4 existe un número natural $K = K(1)$ tal que si $n \geq K$, entonces $\|x_n - x\| \leq 1$. Usando la desigualdad del triángulo se deduce que si $n \geq K$, entonces $\|x_n\| \leq \|x\| + 1$. Si se establece que $M = \sup\{\|x_1\|, \|x_2\|, \ldots, \|x_{K-1}\|, \|x\| + 1\}$, entonces $\|x_n\| \leq$ para toda $n \in \mathbf{N}$.

Se podría pensar que la teoría sobre convergencia de sucesiones en $\mathbf{R}^p$ es más complicada que en $\mathbf{R}$; sin embargo no es este el caso (excepto por cues-

tiones de notación). De hecho el siguiente resultado es importante ya que prueba que los problemas de convergencia en $\boldsymbol{R}^p$ se pueden reducir a problemas idénticos en $\boldsymbol{R}$ para cada una de las sucesiones de coordenadas.

Antes de exponer este resultado recuerde que un elemento típico x de $\boldsymbol{R}^p$ se representa en forma de coordenadas por medio de "p-adas"

$$x = (x_1, x_2, \ldots, x_p).$$

De donde cada elemento de una sucesión (x_n) en $\boldsymbol{R}^p$ tiene una representación similar; por lo que $x_n = (x_{1n}, x_{2n}, \ldots, x_{pn})$. De esta manera la sucesión (x_n) genera p sucesiones de números reales; específicamente $(x_{1n}), (x_{2n}), \ldots, (x_{pn})$. En seguida se demostrará que la convergencia de la sucesión (x_n) se refleja con fidelidad en la convergencia de estas p sucesiones de coordenadas

14.7 TEOREMA. *Una sucesión* (x_n) *en* $\boldsymbol{R}^p$ *con*

$$x_n = (x_{1n}, x_{2n}, \ldots, x_{pn}), \qquad n \in \boldsymbol{N},$$

converge a un elemento $y = (y_1, y_2, \ldots, y_p)$ *si y sólo si las p sucesiones correspondientes de números reales.*

(14.1) $$(x_{1n}), (x_{2n}), \ldots, (x_{pn}),$$

convergen a $y_1, y_2, \ldots, y_p$ *respectivamente.*

DEMOSTRACION. Si $x_n \to y$, entonces $\|x_n - y\| < \varepsilon$ para $n \geq K(\varepsilon)$. Por el teorema 8.10 para cada $j = 1, 2, \ldots, p$, se tiene

$$|x_{jn} - y_j| \leq \|x_n - y\| < \varepsilon, \qquad \text{para} \quad n \geq K(\varepsilon).$$

Por lo tanto cada una de las p sucesiones coordenadas debe converger al número real correspondiente.

De manera inversa suponga que la sucesión en (14.1) converge a y_j para $j = 1, 2, \ldots, p$. Dado $\varepsilon > 0$, hay un número natural $M(\varepsilon)$ al que si $n \geq M(\varepsilon)$, entonces

$$|x_{jn} - y_j| < \varepsilon/\sqrt{p} \qquad \text{para} \quad j = 1, 2, \ldots, p.$$

De aquí se deduce que cuando $n \geq M(\varepsilon)$,

$$\|x_n - y\|^2 = \sum_{j=1}^{p} |x_{jn} - y_j|^2 \leq \varepsilon^2,$$

por lo que la sucesión (x_n) converge a y. Q.E.D.

Ejemplos

En seguida se dan algunos ejemplos estableciéndose la convergencia de una sucesión y usando únicamente los métodos disponibles hasta el momento.

Se observará que para proceder se debe "adivinar" el valor del límite examinando previamente la sucesión. Todos los ejemplos que se presentarán en seguida requieren un poco de "truco" y de habilidad de manipulación; sin embargo los resultados que se obtienen serán muy útiles al establecer (por medio de procedimientos menos artificiales) la convergencia de otras sucesiones. De modo que son de interés tanto los resultados como los métodos.

14.8 EJEMPLOS. *(a)* Sea (x_n) la sucesión en $\boldsymbol{R}$ en donde $x_n = 1/n$. Se demostrará que $\lim (1/n) = 0$. Para hacer esto tome $\varepsilon > 0$; de acuerdo con el corolario 6.7*(b)* (de la propiedad arquimediana) existe un número natural $K(\varepsilon)$ tal que $1/K(\varepsilon) < \varepsilon$. Entonces si $n \geq K(\varepsilon)$ se tiene

$$0 < x_n = \frac{1}{n} \leq \frac{1}{K(\varepsilon)} < \varepsilon,$$

de donde se deduce que $|x_n - 0| < \varepsilon$ para $n \geq K(\varepsilon)$. Dado que $\varepsilon > 0$ es arbitrario esto prueba que $\lim\ (1/n) = 0$.

(b) Sea $a > 0$ y considere la sucesión $X = (1/(1+na))$ en $\boldsymbol{R}$. se habrá de probar que $\lim X = 0$. Observe primero que

$$0 < \frac{1}{1+na} < \frac{1}{na}.$$

Se desea que el término dominante sea menor que una $\varepsilon > 0$ dada para n lo suficientemente grande. De nuevo por el corolario 6.7(*b*) existe un número natural $K(\varepsilon)$ tal que $1/K(\varepsilon) < a\varepsilon$. Entonces, si $n \geq K(\varepsilon)$ se tiene

$$0 < \frac{1}{1+na} < \frac{1}{na} \leq \frac{1}{K(\varepsilon)a} < \varepsilon,$$

por lo que $|1/(1+na) - 0| < \varepsilon$ para $n \geq K(\varepsilon)$. Dado que $\varepsilon > 0$ es arbitrario esto prueba que $\lim X = 0$.

(c) Sea $b \in \boldsymbol{R}$ tal que satisfaga $0 < b < 1$ y considere la sucesión (b^n). Se habrá de probar que $\lim\ (b^n) = 0$ para hacer esto es conveniente escribir b de la siguiente manera

$$b = \frac{1}{1+a}$$

en donde $a > 0$ y usar la desigualdad de Bernoulli $(1+a)^n \geq 1 + na$ para $n \in \boldsymbol{N}$. (Véase el ejercicio 5.C) por lo tanto

$$0 < b^n = \frac{1}{(1+a)^n} \leq \frac{1}{1+na} < \frac{1}{na}.$$

Igual que en el eje anterior está dada si $\varepsilon > 0$ entonces hay un número natural $K(\varepsilon)$ tal que $|b^n - 0| < \varepsilon$ cuando $n \geq K(\varepsilon)$ por lo tanto se tiene $\lim\ (b^n) = 0$.

(d) Sea $c>0$ y considere la sucesión $(c^{1/n})$ se habrá de probar que lim $(c^{1/n})=1$.

Primero suponga que $c>1$. Entonces $c^{1/n}=1+d_n$ con $d_n>0$ de donde, por la desigualdad de Bernoulli

$$c=(1+d_n)^n \geq 1+nd_n.$$

Se deduce que $c-1\geq nd_n$. Dado que $c>1$, se tiene $c-1>0$ por lo tanto dada $\varepsilon>0$, hay un número natural $K(\varepsilon)$ tal que si $n\geq K(\varepsilon)$, entonces

$$0<c^{1/n}-1=d_n\leq\frac{c-1}{n}<\varepsilon.$$

Por lo tanto, $|c^{1/n}-1|<\varepsilon$ cuando $n\geq K(\varepsilon)$, que es lo que se quería.

Ahora suponga que $0<c<1$ (para el caso $c=1$ es obvio). Entonces $c^{1/n}=1/(1+h_n)$ con $h_n>0$ de donde por la desigualdad de Bernoulli,

$$c=\frac{1}{(1+h_n)^n}\leq\frac{1}{1+nh_n}<\frac{1}{nh_n}.$$

Se deduce que $0<h_n<1/nc$ pero como $c>0$, dada $\varepsilon>0$ hay un número natural $K(\varepsilon)$ tal que si $n\geq K(\varepsilon)$ entonces

$$0<1-c^{1/n}=\frac{h_n}{1+h_n}<h_n<\frac{1}{nc}<\varepsilon.$$

Por lo tanto $|c^{1/n}-1|<\varepsilon$ cuando $n\geq K(\varepsilon)$, que es lo que se quería.

(e) Considere la sucesión $X=(n^{1/n})$; se habrá de probar que lim $X=1$, un hecho que no es muy obvio. Escriba $n^{1/n}=1+k_n$ con $k_n>0$ para $n>1$; por lo que $n=(1+k_n)^n$ por el teorema del binomio cuando $n>1$ se tiene

$$n=1+nk_n+\frac{n(n-1)}{2}k_n^2+\cdots>\frac{n(n-1)}{2}k_n^2.$$

Se deduce que $k_n^2<2/(n-1)$, de manera que

$$k_n<\sqrt{\frac{2}{n-1}}.$$

Dada $\varepsilon>0$, existe $K(\varepsilon)$ tal que si $n\geq K(\varepsilon)$, entonces $1/(n-1)<\varepsilon^2/2$; se deduce entonces que $0<k_n<\varepsilon$ y por lo tanto

$$0<n^{1/n}-1=k_n<\varepsilon$$

para $n\geq K(\varepsilon)$. Como $\varepsilon>0$ es arbitraria esto prueba que lim $(n^{1/n})=1$.

Estos ejemplos muestran que resultarán muy útiles ciertos resultados en que no será necesario el ingenio que aquí se ha empleado. Dichos resultados se obtendrán en las dos secciones siguientes. Se cierra esta sección con un resultado que a menudo es útil.

14.9 TEOREMA *Sea $X = (x_n)$ una sucesión en R^p y sea $x \in R^p$. Sea $A = (a_n)$ una sucesión en R tal que*

(i) lim $(a_n) = 0$,

(ii) $\|x_n - x\| \leq C\,|a_n|$ para alguna $C > 0$ y para toda $n \in N$. Entonces, lim $(x_n) = x$.

DEMOSTRACION. Sea $\varepsilon > 0$ dada. Como lim $(a_n) = 0$, existe un número natural $K(\varepsilon)$ tal que si $n \geq K(\varepsilon)$ entonces.

$$C\,|a_n| = C\,|a_n - 0| \leq \varepsilon.$$

Se deduce que

$$\|x_n - x\| \leq C\,|a_n| \leq \varepsilon$$

para toda $n \geq K(\varepsilon)$. Dado que $\varepsilon > 0$ es arbitraria se deduce que lim $(x_n) = x$.

Q.E.D.

Ejercicios

14.A. Sea $b \in R$; demostrar que lim $(b/n) = 0$.

14.B Demostrar que lim $(1/n - 1/(n+1)) = 0$.

14 C. Sea $X = (x_n)$ una sucesión en R^p que converge a x y sea $c \in R$. Demostrar que lim $(cx_n) = cx$.

14.D. Sea $X = (x_n)$ una sucesión en R^p que converge a x. Demostrar que lim $(\|x_n\|) = \|x\|$. (*Sugerencia:* usar la desigualdad del triángulo.)

14.E. Sea $X = (x_n)$ una sucesión en R^p y sea lim $(\|x_n\|) = 0$. Demostrar que lim $(x_n) = 0$. Sin embargo, dar un ejemplo en R para probar que la convergencia de $(|x_n|)$ puede no implicar la convergencia de (x_n).

14.F Demostrar que lim $(1/\sqrt{n}) = 0$. De hecho, si (x_n) es una sucesión de números positivos y lim $(x_n) = 0$, entonces lim $(\sqrt{x_n}) = 0$.

14.G. Sea $d \in R$ tal que $d > 1$. Usar la desigualdad de Bernoulli para demostrar que la sucesión (d^n) no está acotada en R. por lo tanto, no es convergente.

14.H Sea $b \in R$ tal que $0 < b < 1$. Demostrar que lim $(nb^n) = 0$. (*Sugerencia:* usar el teorema del binomio como en el ejemplo 1.8(e).)

14.I Sea $X = (x_n)$ una sucesión de números reales estrictamente positivos tal que lim $(x_{n+1}/x_n) < 1$. Demostrar que para alguna r con $0 < r < 1$ y alguna $C > 0$, se tiene $0 < x_n < Cr^n$ para toda $n \in N$. suficientemente grande. Usar esto para probar que lim $(x_n) = 0$.

14.J. Sea $X = (x_n)$ una sucesión de números reales estrictamente positivos tal que lim $(x_{n+1}/x_n) > 1$. Demostrar que X no es una sucesión acotada y por lo tanto no es convergente.

14.K. Dar un ejemplo de una sucesión convergente (x_n) de números reales estrictamente positivos tal que lim $(x_{n+1}/x_n) = 1$. Dar un ejemplo de una sucesión divergente que tenga esta propiedad

14.L. Aplicar los resultados de los ejercicios 14.I y 14.J a las siguientes sucesiones. (Aquí, $0 < a < 1$, $1 < b$, $c > 0$.)

(a) (a^n),	(b) (na^n),
(c) (b^n),	(d) (b^n/n),
(e) $(c^n/n!)$,	(f) $(2^{3n}/3^{2n})$.

14.M. Sea $X = (x_n)$ una sucesión de números reales estrictamente positivos tal que $\lim (x_n^{1/n}) < 1$. Demostrar que para alguna r con $0 < r < 1$, $0 < x_n < r^n$ para toda $n \in \boldsymbol{N}$ suficientemente grande. Usar esto para deducir que $\lim (x_n) = 0$.

14 N. Sea $X = (x_n)$ una sucesión de números reales estrictamente positivos tal que $\lim (x_n^{1/n}) > 1$. Demostrar que X no es una sucesión acotada y por lo tanto no es convergente.

14.O. Dar un ejemplo de una sucesión convergente (x_n) de números reales estrictamente positivos tal que $\lim (x_n^{1/n}) = 1$. Dar un ejemplo de una sucesión divergente que tenga esta propiedad.

14.P. Reexaminar la convergencia de las sucesiones del ejercicio 14.L a la luz de los ejercicios 14.M y 14.N

14.Q. Analizar la convergencia de las siguientes sucesiones en $\boldsymbol{R}$.

(a) $\left(\frac{(-1)^n}{n}\right)$, (b) $\left(\frac{1}{n^2}\right)$,

(c) $\left(\frac{n^2}{n+1}\right)$. (d) $((-1)^n)$.

Sección 15 Subsucesiones y combinaciones

Esta sección proporciona cierta información acerca de la convergencia de sucesiones que se obtienen de diversas formas a partir de sucesiones que se sabe son convergentes. Esto ayudará a que sea posible extender considerablemente muestra colección de sucesiones convergentes.

15.1 DEFINICION. Si $X = (x_n)$ es una sucesión en $\boldsymbol{R}^p$ y si $r_1 < r_2 < \cdots < r_n < \cdots$ es una sucesión de números naturales extrictamente creciente, entonces la sucesión X' en $\boldsymbol{R}^p$ dada por

$$(x_{r_1}, x_{r_2}, \ldots, x_{r_n}, \ldots),$$

se llama una **subsucesión** de X.

Puede ser útil relacionar el concepto de una subsucesión con el de la composición de dos funciones. Sea g una función, con dominio $\boldsymbol{N}$ y rango en $\boldsymbol{N}$, **estrictamente creciente** en el sentido de que si $n < m$; entonces $g(n) < g(m)$. De donde g define una subsucesión de $X = (x_n)$ por medio de la fórmula

$$X \circ g = (x_{g(n)} : n \in \boldsymbol{N}).$$

A la inversa toda subsucesión de X es de la forma $X \circ g$ para alguna función estrictamente creciente con $D(g) = \boldsymbol{N}$ y $R(g) \subseteq \boldsymbol{N}$.

Es claro que una sucesión dada tiene muchas subsucesiones distintas. A pesar de que el siguiente resultado es muy elemental, es de suficiente importancia como para hacerlo explícito.

15.2 LEMA. *Si alguna sucesión X en $\mathbf{R}^p$ converge a un elemento x, entonces cualquier subsucesión de X también converge a x.*

DEMOSTRACION. Sea V una vecindad del elemento límite x; por definición, existe un número natural K_V tal que para toda $n \geq K_V$, x_n pertenece a V. Ahora, sea X' una subsucesión de X; digamos

$$X' = (x_{r_1}, x_{r_2}, \ldots, x_{r_n}, \ldots).$$

Dado que $r_n \geq n$, entonces $r_n \geq K_V$ y por lo tanto x_{r_n} pertenece a V. Esto demuestra que X' también convergente a x. Q.E.D.

15.3 COROLARIO *Si $X = (x_n)$ es una sucesión que converge a un elemento x de $\mathbf{R}^p$ y si m es cualquier número natural, entonces la sucesión $X' = (x_{m+1}, x_{m+2}, \ldots)$ también converge a x.*

DEMOSTRACION. Dado que X' es una subsucesión de X, el resultado se deduce directamente del lema anterior. Q.E.D.

Los últimos resultados han sido orientados de manera especial hacia la demostración de que una sucesión converge a un punto dado. También es importante saber lo que significa exactamente que una sucesión X *no* converja a x. El siguiente resultado es elemental pero no trivial y su comprobación desempeña un papel importante en la preparación del lector. Por lo tanto, se le deja la demostración detallada.

15.4 TEOREMA *Si $X = (x_n)$ es una sucesión en $\mathbf{R}^p$, entonces las siguientes afirmaciones son equivalentes:*

(a) X no converge a x.

(b) Existe una vecindad V de x tal que si n es cualquier número natural entonces hay un número natural $m = m(n) \geq n$ tal que x_m no pertenece a V.

(c) Existe una vecindad V de x y una subsucesión X' de X tal que ninguno de los elementos de X' pertenece a V.

15.5 EJEMPLOS. *(a)* Sea X la sucesión en $\mathbf{R}$ que consta de los números naturales.

$$X = (1, 2, \ldots, n, \ldots).$$

Sea x cualquier número real y considere la vecindad V de x que consta del intervalo abierto $(x-1, x+1)$. acuerdo con la propiedad arquimediana 6.6 existe un número natural k_0 tal que $x+1 < k_0$; de donde, si $n \geq k_0$, se deduce que $x_n = n$ no pertenece a V por lo tanto, la subsucesión $X' = (k_0, k_0+1, \ldots)$ X no tiene puntos en V, probándose que X no converge a x.

(b) Sea $Y = (y_n)$ la sucesión en $\mathbf{R}$ que consta de $Y = (-1, 1, \ldots, (-1)^n, \ldots)$. Se deja al lector la demostración de que ningún punto y, excepto posiblemente $y = \pm 1$, puede ser un límite de Y. Se probará que el punto $y = -1$

no es un límite de Y; la demostración para $y = +1$ es por completo análoga. Sea V la vecindad de $y = -1$ que consta del intervalo abierto $(-2, 0)$. Entonces, si n es par, el elemento $y_n = (-1)^n = +1$ no pertenece a V. Por lo tanto, la subsucesión Y' de Y que corresponde a $r_n = 2n$, $n \in \mathbf{N}$, evade la vecindad V, probándose que $y = -1$ no es límite de Y.

(c) Sea $Z = (z_n)$ una sucesión en $\mathbf{R}$ con $z_n \geq 0$, para $n \geq 1$. Se deduce que ningún número $z < 0$ puede ser límite de Z. De hecho, el intervalo abierto $V = \{x \in \mathbf{R} : x < 0\}$ es una vecindad de z que no contiene ningún elemento Z. Esto prueba (¿por qué?) que z no puede ser el límite de Z. por lo tanto, si Z tiene un límite, este límite debe ser positivo.

Combinaciones de sucesiones

El siguiente teorema permite usar las operaciones algebraicas de las definiciones 14.2 para formar nuevas sucesiones cuya convergencia se puede predecir a partir de la convergencia de las sucesiones dadas.

15.6 TEOREMA. *(a) Sean X y Y sucesiones en $\mathbf{R}^p$ que convergen en x y y, respectivamente. Entonces, las sucesiones $X + Y$, $X - Y$, y $X \cdot Y$ convergen a $x + y$, $x - y$, y $x \cdot y$, respectivamente.*

(b) Sea $X = (x_n)$ una sucesión en $\mathbf{R}^p$ que converge a x y sea $A = (a_n)$ una sucesión en $\mathbf{R}$ que converge a a. Entonces, la sucesión $(a_n x_n)$ en $\mathbf{R}^p$ converge a ax.

(c) Sea $X = (x_n)$ una sucesión en $\mathbf{R}^p$ que converge a x y sea $B = (b_n)$ una sucesión de números reales distintos de cero que converge a un número distinto de cero b. Entonces, la sucesión $(b_n^{-1} x_n)$ en $\mathbf{R}^p$ converge a $b^{-1}x$.

DEMOSTRACION. *(a)* Para probar que $(x_n + y_n) \to x + y$, es necesario calcular la magnitud de $\|(x_n + y_n) - (x + y)\|$. Para hacer esto, se usa la desigualdad del triángulo y se obtiene

$$\begin{aligned} \|(x_n + y_n) - (x + y)\| &= \|(x_n - x) + (y_n - y)\| \\ &\leq \|x_n - x\| + \|y_n - y\|. \end{aligned} \tag{15.1}$$

Por hipótesis, si $\varepsilon > 0$, se puede elegir K_1 tal que si $n \geq K_1$, entonces $\|x_n - x\| < \varepsilon/2$ y se elige K_2 tal que si $n \geq K_2$, entonces $\|y_n - y\| < \varepsilon/2$. Por lo tanto, si $K_0 = \sup\{K_1, K_2\}$ y $n \geq K_0$, de (15.1) se concluye que

$$\|(x_n + y_n) - (x + y)\| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Dado que esto se puede hacer para cualquier $\varepsilon > 0$, arbitraria, se deduce que $X + Y$ converge a $x + y$. Exactamente el mismo argumento se puede usar para probar que $X - Y$ converge a $x - y$.

Para demostrar que $X \cdot Y$ converge a $x \cdot y$ se hace el cálculo

$$\begin{aligned} |x_n \cdot y_n - x \cdot y| &= |(x_n \cdot y_n - x_n \cdot y) + (x_n \cdot y - x \cdot y)| \\ &\leq |x_n \cdot (y_n - y)| + |(x_n - x) \cdot y|. \end{aligned}$$

Usando la desigualdad de Schwarz, se obtiene

$$(15.2) \qquad |x_n \cdot y_n - x \cdot y| \le \|x_n\| \, \|y_n - y\| + \|x_n - x\| \, \|y\|.$$

De acuerdo con el lema 14.6, existe un número $M > 0$ que es cota superior de $\{\|x_n\|, \|y\|\}$. Además, por la convergencia de X, Y se concluye que si está dada $\varepsilon > 0$ entonces existen números naturales K_1, K_2 tales que si $n \ge K_1$, entonces $\|y_n - y\| < \varepsilon/2M$ y si $n \ge K_2$, entonces $\|x_n - x\| < \varepsilon/2M$. Ahora, elíjase $K = \sup\{K_1, K_2\}$; entonces, si $n \ge K$, de (15.2) se deduce que

$$|x_n \cdot y_n - x \cdot y| \le M \|y_n - y\| + M \|x_n - x\|$$

$$< M\left(\frac{\varepsilon}{2M} + \frac{\varepsilon}{2M}\right) = \varepsilon.$$

Esto prueba que $X \cdot Y$ converge a $x \cdot y$.

La parte *(b)* se demuestra de la misma manera.

Para demostrar *(c)* se hacen los siguientes cálculos:

$$\left\|\frac{1}{b_n} x_n - \frac{1}{b} x\right\| = \left\|\left(\frac{1}{b_n} x_n - \frac{1}{b} x_n\right) + \left(\frac{1}{b} x_n - \frac{1}{b} x\right)\right\|$$

$$\le \left|\frac{1}{b_n} - \frac{1}{b}\right| \|x_n\| + \frac{1}{|b|} \|x_n - x\|$$

$$= \frac{|b - b_n|}{|b_n b|} \|x_n\| + \frac{1}{|b|} \|x_n - x\|.$$

Ahora, sea $M > 0$ tal que

$$\frac{1}{M} < |b| \qquad \text{y} \qquad \|x\| < M.$$

Se deduce que existe un número natural K_0 tal que si $n \ge K_0$, entonces

$$\frac{1}{M} < |b_n| \qquad \text{y} \qquad \|x_n\| < M.$$

De donde, si $n \ge K_0$, del cálculo anterior se obtiene

$$\left\|\frac{1}{b_n} x_n - \frac{1}{b} x\right\| \le M^3 |b_n - b| + M \|x_n - x\|.$$

Por lo tanto, si $\varepsilon > 0$ es un número real ya fijado, entonces existen números naturales K_1, K_2 tales que si $n \ge K_1$ entonces $|b_n - b| < \varepsilon/2M^3$ y si $n \ge K_2$ entonces $\|x_n - x\| < \varepsilon/2M$. Siendo $K = \sup\{K_0, K_1, K_2\}$ se concluye que si $n \ge K$, entonces

$$\left\|\frac{1}{b_n} x_n - \frac{1}{b} x\right\| < M^3 \frac{\varepsilon}{2M^3} + M \frac{\varepsilon}{2M} = \varepsilon,$$

lo cual prueba que (x_n/b_n) converge a x/b. Q.E.D.

15.7 APLICACIONES. De nuevo se limitará la atención a sucesiones en $\boldsymbol{R}$.

(a) Sea $X = (x_n)$ la sucesión en $\boldsymbol{R}$ definida por

$$x_n = \frac{2n+1}{n+5}, \qquad n \in \boldsymbol{N}.$$

Observe que se puede escribir x_n de la siguiente manera

$$x_n = \frac{2+1/n}{1+5/n};$$

de modo que X se puede considerar como el cociente de $Y = (2+1/n)$ y $Z = (1+5/n)$. Dado que la última sucesión consta de términos distintos de cero y tiene límite 1 (¿por qué?), el teorema anterior es aplicable y se puede concluir que

$$\lim X = \frac{\lim Y}{\lim Z} = \frac{2}{1} = 2.$$

(b) Si $X = (x_n)$ es una sucesión en $\boldsymbol{R}$ que converge a x y si p es un polinomio, entonces la sucesión definida por $(p(x_n) : n \in \boldsymbol{N})$ converge a $p(x)$. (Sugerencia: usar el teorema 15.6 e inducción.)

(c) Sea $X = (x_n)$ una sucesión en $\boldsymbol{R}$ que converge a x y sea r una función racional; es decir, $r(y) = p(y)/q(y)$, en donde p y q son polinomios. Suponga que $q(x_n)$ y $q(x)$ son distintos de cero, entonces la sucesión $(r(x_n) : n \in \boldsymbol{N})$ converge a $r(x)$. (*Sugerencia:* usar la parte *(b)* y el teorema 15.6.)

Se concluye esta sección con un resultado que a menudo es útil. Algunas veces se describe diciendo que "se alcanza el límite en una desigualdad".

15.8 LEMA *Suponga que $X = (x_n)$ es una sucesión convergente en $\boldsymbol{R}^p$ con límite x. Si existen algún elemento c en $\boldsymbol{R}^p$ y un número $r > 0$ tal que $\|x_n - c\| \le r$ para n suficientemente grande, entonces $\|x - c\| \le r$.*

DEMOSTRACION. El conjunto $V = \{y \in \boldsymbol{R}^p : \|y - c\| > r\}$ es un subconjunto abierto de $\boldsymbol{R}^p$. Si $x \in V$, entonces V es una vecindad de x y $x_n \in V$ para valores de n lo suficientemente grandes, contrario a la hipótesis. Por lo tanto, $x \notin V$ y se tiene $\|x - c\| \le r$. Q.E.D.

Es importante observar que en este resultado se ha supuesto la existencia del límite, ya que las hipótesis restantes no son suficientes para demostrar su existencia.

Ejercicios

15.A. Si (x_n) y (y_n) son sucesiones convergentes de números reales y si $x_n \le y_n$ para toda $n \in \boldsymbol{N}$, entonces $\lim (x_n) \le \lim (y_n)$.

15.B. Si $X = (x_n)$ y $Y = (y_n)$ son sucesiones de números reales que convergen a c y si $Z = (z_n)$ es una sucesión tal que $x_n \le z_n \le y_n$ para $n \in \boldsymbol{N}$, entonces también converge a c.

15.C. para x_n dada mediante las siguientes fórmulas, establecer la convergencia o la divergencia de la sucesión $X = (x_n)$:

(a) $x_n = \dfrac{n}{n+1}$, (b) $x_n = \dfrac{(-1)^n n}{n+1}$,

(c) $x_n = \dfrac{2n}{3n^2+1}$, (d) $x_n = \dfrac{2n^2+3}{3n^2+1}$,

(e) $x_n = n^2 - n$, (f) $x_n = \text{sen } n$.

15.D. Si X y Y son sucesiones en $\boldsymbol{R}^p$ y si $X + Y$ converge, ¿convergen X y Y y tienen límite $\lim (X+Y) = \lim X + \lim Y$?

15.E. Si X y Y son sucesiones en $\boldsymbol{R}^p$ y si $X \cdot Y$ converge, ¿convergen X y Y y tienen $\lim X \cdot Y = (\lim X) \cdot (\lim Y)$?

15.F. Si $X = (x_n)$ es una sucesión positiva que converge a x, entonces $(\sqrt{x_n})$ converge a $\sqrt{x}$. (*Sugerencia:* $\sqrt{x_n} - \sqrt{x} = (x_n - x)/(\sqrt{x_n} + \sqrt{x})$ cuando $x \neq 0$.)

15.G. Si $X = (x_n)$ es una sucesión de números reales tal que $Y = (x_n^2)$ converge a 0, ¿converge X a 0?

15.H. Si $x_n = \sqrt{n+1} - \sqrt{n}$, ¿convergen las sucesiones $X = (x_n)$ y $Y = (\sqrt{n}\, x_n)$?

15.I. Sea (x_n) una sucesión en $\boldsymbol{R}^p$ tal que las subsucesiones (x_{2n}) y (x_{2n+1}) convergen a $x \in \boldsymbol{R}^p$. Demostrar que (x_n) converge a x.

15.J. Sean (x_n) y (y_n) sucesiones en $\boldsymbol{R}$ tales que $\lim (x_n) \neq 0$ y $\lim (x_n y_n)$ existe. Demostrar que $\lim (y_n)$ existe

15.K. ¿Sigue siendo válido el ejercicio 15.J en $\boldsymbol{R}^2$?

15.L. Si $0 < a \leq b$ y si $x_n = (a^n + b^n)^{1/n}$, entonces $\lim (x_n) = b$.

15.M. Todo número irracional en $\boldsymbol{R}$ es el límite de una sucesión de números racionales. Todo número racional en $\boldsymbol{R}$ es el límite de una sucesión de números irracionales.

15.N. Sean $A \subseteq \boldsymbol{R}^p$ y $x \in \boldsymbol{R}^p$. Entonces x es un punto frontera de A si y sólo si hay una sucesión (a_n) de elementos en A y una sucesión (b_n) de elementos en $\mathscr{C}(A)$ tales que

$$\lim (a_n) = x = \lim (b_n).$$

15.O. Sea $A \subseteq \boldsymbol{R}^p$ y $x \in \boldsymbol{R}^p$. Entonces, x es un punto de acumulación de A si y sólo si hay una sucesión (a_n) de elementos distintos en A tal que $x = \lim (a_n)$.

15.P. $x = \lim (x_n)$ y si $\|x_n - c\| < r$ para toda $n \in \boldsymbol{N}$, ¿se deduce que $\|x - c\| < r$?

Proyectos

15.α. Sea d una métrica en un conjunto M en el sentido que se da en el ejercicio 8.S. Si $X = (x_n)$ es una sucesión en M, entonces un elemento $x \in M$ se dice que es un **límite** de X si para cada $\varepsilon > 0$ existe un número $K(\varepsilon)$ en N tal que para toda $n \geq K(\varepsilon)$, $d(x_n, x) < \varepsilon$. Usar esta definición y demostrar que los teoremas 14.5, 14.4, 15.2, 15.3 y 15.4 se pueden extender a espacios métricos. Demostrar que las métricas d_1, d_2, d_∞ en $\boldsymbol{R}^p$ dan origen a las mismas sucesiones convergentes en $\boldsymbol{R}^p$. Demostrar que si d es una métrica discreta en un conjunto entonces las únicas sucesiones que convergen con respecto a d son aquellas que son "constantes respecto de algún número natural".

15.β. Denótese por medio de m a la colección de todas las sucesiones acotadas en $\boldsymbol{R}$, por medio de c a la colección de todas las sucesiones convergentes en $\boldsymbol{R}$ y por medio de c_0 a la colección de todas las sucesiones en $\boldsymbol{R}$ que convergen a cero.

(a) Con la suma $X+Y$ y el producto cX de la definición 14.2, demostrar que cada una de las colecciones anteriores es un espacio vectorial en donde el elemento cero es la sucesión $0=(0,0,\ldots)$.

(b) En cada una de las colecciones m, c, c_0, definir la norma de $X=(x_n)$ como $\|x\|=\sup\{|x_n|: n\in \boldsymbol{N}\}$. Demostrar que esta definición realmente da una norma.

(c) Si X y Y pertenecen a m, a c, o a c_0, entonces el producto XY también pertenece a la colección y $\|XY\|\le\|X\|\,\|Y\|$. Dar un ejemplo para probar que la igualdad puede ser válida en esta última relación y otro para probar que puede no cumplirse la igualdad.

(d) Demostrar que la métrica inducida por la norma, en la parte *(b)*, en estos espacios está dada por $d(X,Y)=\sup\{|x_n-y_n|: n\in \boldsymbol{N}\}$.

(e) Demostrar que si una sucesión (X_k) converge a Y con respecto a la en *(d)* entonces cada "sucesión coordenada" converge a la coordenada correspondiente de Y. (Observación: X_k es una sucesión en $\boldsymbol{R}$, mientras que (X_k) es una sucesión en m, c o c_0; es decir, una "sucesión de sucesiones" en $\boldsymbol{R}$.)

(f) Dar un ejemplo de una sucesión (X_k) en c_0 en donde cada sucesión coordenada converja a 0, pero en donde $d(X_k, 0)$ no converja a 0.

Sección 16 Dos criterios de convergencia

Hasta ahora, el método principal disponible para probar que una sucesión es convergente es el de identificarla como una subsucesión de una combinación algebraica de sucesiones convergentes. Cuando es posible hacer esto, se puede calcular el límite usando los resultados de la sección anterior. Sin embargo, cuando esto no se puede hacer, es necesario apoyarse en la definición 14.3 o en el teorema 14.4 para poder establecer la existencia del límite. El uso de estos últimos medios tiene la notable desventaja de que se debe saber de antemano (o cuando menos suponer) el valor correcto del límite y después verificar el supuesto resultado.

Sin embargo, existen muchos casos en que no hay ningún candidato obvio para el límite de una sucesión dada, aun cuando un análisis preliminar haga pensar que sí hay convergencia. En esta sección se dan algunos resultados más a fondo que los de las secciones anteriores y que se pueden usar para establecer la convergencia de una sucesión cuando ningún elemento en particular aparece como el valor del límite. El primer resultado con este enfoque es muy importante. A pesar de que se puede generalizar a $\boldsymbol{R}^p$, es conveniente restringir el enunciado para el caso de sucesiones en $\boldsymbol{R}$.

16.1 TEOREMA DE CONVERGENCIA MONOTONA. *Sea $X=(x_n)$ una sucesión de números reales monótonamente creciente en el sentido de que*

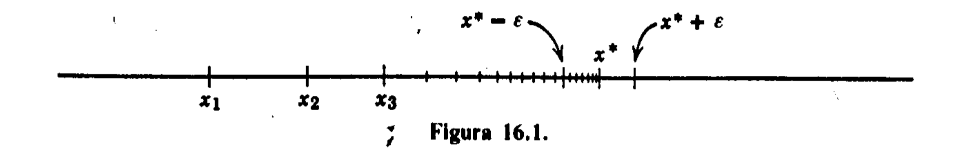


Figura 16.1.

$$x_1 \leq x_2 \leq \cdots \leq x_n \leq x_{n+1} \leq \cdots.$$

Entonces, la sucesión X converge si y sólo si es acotada, en cuyo caso

$$\lim (x_n) = \sup \{x_n\}.$$

DEMOSTRACION. En el lema 14.6 se vio que una sucesión convergente es acotada. Si $x = \lim (x_n)$ y $\varepsilon > 0$, entonces existe un número natural $K(\varepsilon)$ tal que si $n \geq K(\varepsilon)$, entonces

$$x - \varepsilon \leq x_n \leq x + \varepsilon.$$

Dado que X es monótona, de esta relación se obtiene

$$x - \varepsilon \leq \sup \{x_n\} \leq x + \varepsilon,$$

y se infiere que $|x - \sup \{x_n\}| \leq \varepsilon$. Como esto es válido para toda $\varepsilon > 0$, se deduce que $\lim (x_n) = x = \sup \{x_n\}$.

A la inversa suponga que $X = (x_n)$ es una sucesión acotada de números reales monótonamente creciente. De acuerdo con el principio del supremo, el supremo $x^* = \sup \{x_n\}$ existe; se probará que es el límite de X. Dado que x^* es una cota superior de los elementos en X, entonces $x_n \leq x^*$ para $n \in \mathbf{N}$. Como x^* es el supremo de Z, si $\varepsilon > 0$ el número $x^* - \varepsilon$ no es cota superior de X y existe un número natural $K(\varepsilon)$ tal que

$$x^* - \varepsilon < x_{K(\varepsilon)}.$$

Debido a la propiedad de monotonía de X, para toda $n \geq K(\varepsilon)$,

$$x^* - \varepsilon < x_n \leq x^*,$$

por lo que se deduce que $|x_n - x^*| < \varepsilon$. Resumiendo lo anterior, el número $x^* = \sup \{x_n\}$ tiene la propiedad de que, dado $\varepsilon > 0$, hay un número natural $K(\varepsilon)$ (que depende de ε) tal que $|x_n - x^*| < \varepsilon$ siempre que $n \geq K(\varepsilon)$. Esto prueba que $x^* = \lim X$. Q.E.D.

16.2 COROLARIO. *Sea $X = (x_n)$ una sucesión de números reales monótonamente decreciente en el sentido de que*

$$x_1 \geq x_2 \geq \cdots \geq x_n \geq x_{n+1} \geq \cdots.$$

Entonces, la sucesión X converge si y sólo si es acotada, en cuyo caso

$$\lim (x_n) = \inf \{x_n\}.$$

DEMOSTRACION. Sea $y_n = -x_n$ para $n \in \mathbf{N}$. Entonces se puede ver que con facilidad la sucesión $Y = (y_n)$ es una sucesión monótonamente cre-

ciente. Más aún, Y es acotada si y sólo si X es acotada. Por lo tanto, la conclusión se deduce del teorema. Q.E.D.

16.3 EJEMPLOS. *(a)* Volviendo a la sucesión $X = (1/n)$ que se vio en el ejemplo 14.8(*a*), es claro que

$$\frac{1}{1} > \frac{1}{2} > \cdots > \frac{1}{n} > \cdots > 0;$$

entonces del corolario 16.2 se deduce que $X = (1/n)$ converge. Se puede establecer el valor de lim $(1/n)$ siempre y cuando se pueda calcular inf $\{1/n\}$. Alternativamente, una vez asegurada la convergencia de X, a menudo se puede calcular su límite usando el lema 15.2 y el teorema 15.6. En el caso que se está considerando, si $X' = (1/2, 1/4, \ldots, 1/2n, \ldots)$, entonces se deduce que

$$\lim X = \lim X' = \tfrac{1}{2} \lim X.$$

por lo tanto, se concluye que $\lim X = 0$.

(b) Sea $Y = (y_n)$ una sucesión en $\boldsymbol{R}$ definida inductivamente por

$$y_1 = 1, \qquad y_{n+1} = (2y_n + 3)/4 \qquad \text{para} \quad n \in \boldsymbol{N}.$$

por cálculo directo se tiene $y_1 < y_2 < 2$. Si $y_{n-1} < y_n < 2$, entonces

$$2y_{n-1} + 3 < 2y_n + 3 < 2 \cdot 2 + 3,$$

de donde se infiere que $y_n < y_{n+1} < 2$ por inducción, la sucesión Y es monótonamente creciente y está acotada por arriba con el número 2. Del teorema de convergencia monótona se infiere que la sucesión Y converge a un límite que no es mayor a 2. En este caso podría no ser tan fácil evaluar $y = \lim Y$ calculando sup $\{y_n\}$. Sin embargo, sabiendo ya que el límite existe, hay otra manera de calcular su valor. De acuerdo con el lema 15.2 se tiene $y = \lim (y_n) = \lim (y_{n+1})$. Usando el teorema 15.6, el límite y debe satisfacer la relación

$$y = (2y + 3)/4.$$

Por lo tanto, se concluye que $y = \frac{3}{2}$.

(c) Sea $Z = (z_n)$ una sucesión en $\boldsymbol{R}$ definida por

$$z_1 = 1, \qquad z_{n+1} = \sqrt{2z_n} \qquad \text{para} \quad n \in \boldsymbol{N}.$$

Es claro que $z_1 < z_2 < 2$. Si $z_n < z_{n+1} < 2$, entonces $2z_n < 2z_{n+1} < 4$ de modo que $z_{n+1} = \sqrt{2z_n} < z_{n+2} = \sqrt{2z_{n+1}} < 2 = \sqrt{4}$. Esto prueba que Z es una sucesión monótonamente creciente acotada por arriba por 2, por lo que Z converge a un número z. Se puede demostrar en forma directa que $2 = \sup \{z_n\}$ de manera que el límite $z = 2$. De otro modo, se puede usar el método del ejemplo anterior. Sabiendo que la sucesión tiene un límite z, se concluye de la relación $z_{n+1} = \sqrt{2z_n}$ que z debe satisfacer $z = \sqrt{2z}$. Para encontrar las raíces de esta última ecuación se eleva al cuadrado para obtener $z^2 = 2z$, que tiene raíces

0,2. Es evidente que 0 no puede ser el límite (¿por qué?); por lo tanto, este límite debe ser igual a 2.

(d) Sea $U=(u_n)$ una sucesión de números reales definida por $u_n=(1+1/n)^n$ para $n \in \mathbf{N}$. Aplicando el teorema del binomio se puede escribir

$$u_n = 1+\frac{n}{1}\frac{1}{n}+\frac{n(n-1)}{2!}\frac{1}{n^2}+\frac{n(n-1)(n-2)}{3!}\frac{1}{n^3}$$
$$+\cdots+\frac{n(n-1)\cdots 2\cdot 1}{n!}\frac{1}{n^n}.$$

Dividiendo las potencias de n entre los numeradores de los coeficientes binomiales se tiene

$$u_n = 1+1+\frac{1}{2!}\left(1-\frac{1}{n}\right)+\frac{1}{3!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)$$
$$+\cdots+\frac{1}{n!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\cdots\left(1-\frac{n-1}{n}\right).$$

Expresando u_{n+1} de la misma manera se tiene

$$u_{n+1} = 1+1+\frac{1}{2!}\left(1-\frac{1}{n+1}\right)+\frac{1}{3!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)$$
$$+\cdots+\frac{1}{n!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\cdots\left(1-\frac{n-1}{n+1}\right)$$
$$+\frac{1}{(n+1)!}\left(1-\frac{1}{n+1}\right)\left(1-\frac{2}{n+1}\right)\cdots\left(1-\frac{n}{n+1}\right).$$

Observe que la expresión de u_n contiene $n + 1$ términos y la de u_{n+1} contiene $n + 2$ términos. Mediante un análisis elemental se prueba que cada término en u_n nunca es mayor al término correspondiente en u_{n+1} que este tiene un término positivo de más. por lo tanto, se tiene

$$u_1 < u_2 < \cdots < u_n < u_{n+1} < \cdots.$$

Para probar que la sucesión está acotada, observe que si $p=1, 2, \ldots, n,$ entonces $(1-p/n)<1$. Más aún, $2^{p-1} \leq p!$ (¿por qué?) de manera que $1/p! \leq 1/2^{p-1}$. De la expresión anterior para u_n y a partir de estos cálculos se obtiene

$$2 < u_n < 1+1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^{n-1}} < 3, \qquad n>2.$$

Se deduce que la sucesión monótona U está acotada en 3 por arriba. El teorema de convergencia monótona implica que la sucesión U converge a un número real que cuando mucho es igual a 3. Recuerde que el límite de U es el número fundamental e. Si se perfeccionan los cálculos se puede encontrar una aproximación racional más cercana al valor de e; pero no se puede evaluar exactamente de esta manera ya que es irracional. Sin embargo, es posible calcular todas las cifras decimales que se deseen. Esto pone en claro que no es un resultado como el teorema de convergencia monótona, que sólo establece la existencia del límite de una sucesión, puede ser de gran utilidad aun cuando el valor exacto no se pueda obtener con facilidad.

El teorema de Bolzano-Weierstrass

El teorema de convergencia monótona es extraordinariamente útil e importante, pero tiene el inconveniente de que sólo se aplica a sucesiones monótonas. Nos corresponde, entonces, encontrar una condición que implique convergencia en $\boldsymbol{R}$ o $\boldsymbol{R}^p$ sin usar la propiedad de monotonía. Esta condición deseada es el criterio de Cauchy que se dará en seguida. Sin embargo, primero se dará una forma del teorema de Bolzano-Weierstrass 10.6 que es aplicable en especial a sucesiones.

16.4 TEOREMA DE BOLZANO-WEIERSTRASS. *Una sucesión acotada en $\boldsymbol{R}^p$ tiene una subsucesión convergente.*

DEMOSTRACION. Sea $X = (x_n)$ una sucesión acotada en $\boldsymbol{R}^p$. Si sólo hay un número finito de valores distintos en la sucesión X, entonces cuando menos uno de estos valores debe aparecer un número infinito de veces. Si se define una su sucesión de X, seleccionando este elemento cada vez que aparece, se obtiene una subsucesión convergente de X.

Por otro lado, si la sucesión X contiene un número infinito de valores distintos en $\boldsymbol{R}^p$, entonces, puesto que estos puntos están acotados, el teorema de Bolzano-Weierstrass 10.6 para conjuntos implica que hay cuando menos un punto de acumulación, digamos x^*. Sea x_{n_1} un elemento de X tal que

$$\|x_{n_1} - x^*\| < 1.$$

Considere la vecindad $V_2 = \{y : \|y - x^*\| < \frac{1}{2}\}$. Dado que el punto x^* es un punto de acumulación del conjunto $S_1 = \{x_m : m \geq 1\}$, también es un punto de aculación del conjunto $S_2 = \{x_m : m > n_1\}$ que se obtiene al suprimir un número finito de elementos de S_1 (¿por qué?) Por lo tanto, hay un elemento x_{n_2} de S_2 (por lo que $n_2 > n_1$) que pertenece a V_2. Ahora, sea V_3 la vecindad $V_3 = \{y : \|y - x^*\| < \frac{1}{3}\}$ y sea $S_3 = \{x_m : m > n_2\}$. Dado que x^* es un punto de acumulación de S_3 dente haber un elemento x_{n_3} de S_3 (por lo que $n_3 > n_2$) que pertenezca a V_3. Continuando de esta manera se obtiene una subsucesión $X' = (x_{n_1}, x_{n_2}, \ldots)$ de X con

$$\|x_{n_r} - x^*\| < 1/r,$$

por lo que $\lim X' = x^*$. **Q.E.D.**

16.5 COROLARIO. *Si $X = (x_n)$ es una sucesión en $\boldsymbol{R}^p$ y x^* es un punto de acumulación del conjunto $\{x_n : n \in \mathbf{N}\}$, entonces existe una subsucesión X' de X que converge a x^*.*
De hecho, esto es lo que establece la segunda parte de la demostración de 16.4.

Sucesiones de Cauchy

En seguida se da el concepto tan importante de una sucesión de Cauchy en $\mathbf{R}^p$. Resultará que una sucesión en $\mathbf{R}^p$ es convergente si y sólo si es una sucesión de Cauchy.

16.6 DEFINICION. Se dice que una sucesión $X = (x_n)$ en $\mathbf{R}^p$ es una **sucesión de Cauchy** si para toda $\varepsilon > 0$ hay un número natural $M(\varepsilon)$ tal que para todas $m, n \geq M(\varepsilon)$, $\|x_m - x_n\| < \varepsilon$.

Para ayudar a aclarar el concepto de una sucesión de Cauchy se demostrará que toda sucesión convergente en $\mathbf{R}^p$ es una sucesión de Cauchy.

16.7 LEMA *Si $X = (x_n)$ es una sucesión convergente en $\mathbf{R}^p$, entonces X es una sucesión de Cauchy.*

DEMOSTRACION. $x = \lim X$; entonces, dada $\varepsilon > 0$ hay un número natural $K(\varepsilon/2)$ tal que si $n \geq K(\varepsilon/2)$, entonces $\|x_n - x\| < \varepsilon/2$. De modo que si $M(\varepsilon) = K(\varepsilon/2)$ y si $m, n \geq M(\varepsilon)$, entonces

$$\|x_m - x_n\| \leq \|x_m - x\| + \|x - x_n\| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Por lo que la sucesión convergente X es una sucesión de Cauchy. Q.E.D.

Para poder aplicar el teorema de Bolzano-Weierstrass es necesario el siguiente resultado.

16.8 LEMA. *Una sucesión de Cauchy en $\mathbf{R}^p$ es acotada.*

DEMOSTRACION. Sea $X = (x_n)$ una sucesión de Cauchy y sea $\varepsilon = 1$. Si $m = M(1)$ y $n \geq M(1)$, entonces $\|x_m - x_n\| < 1$. Por la desigualdad del triángulo esto implica que $\|x_n\| < \|x_m\| + 1$ para $n \geq M(1)$. Por lo tanto, si

$$B = \sup\{\|x_1\|, \ldots, \|x_{m-1}\|, \|x_m\| + 1\},$$

se tiene $\|x_n\| \leq B$ para toda $n \in \mathbf{N}$. Por lo tanto, la sucesión de Cauchy X es acotada. Q.E.D.

16.9 LEMA. *Si una subsucesión X' de una sucesión de Cauchy X en $\mathbf{R}^p$ converge a un elemento x, entonces toda la sucesión X converge a x*

DEMOSTRACION. Puesto que $X = (x_n)$ es una sucesión de Cauchy, dada $\varepsilon > 0$ hay un número natural $M(\varepsilon/2)$ tal que si $m, n \geq M(\varepsilon/2)$,

$$(*) \qquad \|x_m - x_n\| < \varepsilon/2.$$

Si la sucesión $X' = (x_{n_j})$ converge a x, hay un número natural $K \geq M(\varepsilon/2)$, que pertenece al conjunto $\{n_1, n_2, \ldots\}$ y tal que

$$\|x - x_K\| < \varepsilon/2.$$

Ahora, sea n cualquier número natural tal que $n \geq M(\varepsilon/2)$. Se deduce que (*) e válida para este valor de n y para . $m = K$. De modo que

$$\|x - x_n\| \leq \|x - x_K\| + \|x_K - x_n\| < \varepsilon,$$

cuando $n \geq M(\varepsilon/2)$. Por lo tanto, la sucesión X converge al elemento x, que es el límite de la subsucesión X'. Q.E.D.

Ahora, ya es posible obtener el importante criterio de Cauchy. La demostración es engañosamente corta, el lector se podrá dar cuenta de que el trabajo ya se ha hecho y que sólo se están juntando las distintas partes.

16.10 CRITERIO DE CONVERGENCIA DE CAUCHY. *Una sucesión en* $\boldsymbol{R}^p$ *es convergente si y sólo si es una sucesión de Cauchy.*

DEMOSTRACION. Se vio en el lema 16.7 que una sucesión convergente debe ser una sucesión de Cauchy.

A la inversa, suponga que X es una sucesión de Cauchy en $\boldsymbol{R}^p$. Del lema 16.8 se deduce que la sucesión X está acotada en $\boldsymbol{R}^p$. De acuerdo con el teorema de Bolzano-Weierstrass 16.4 la sucesión acotada X tiene una subsucesión convergente X'. Por el lema 16.9 toda la sucesión X converge al límite de X'.

16.11 EJEMPLOS. *(a)* Sea $X = (x_n)$ una sucesión en R definida por

$$x_1 = 1, x_2 = 2, \ldots, x_n = \tfrac{1}{2}(x_{n-2} + x_{n-1}) \quad \text{para} \quad n > 2.$$

Se puede demostrar por inducción que

$$1 \leq x_n \leq 2 \quad \text{para} \quad n \in \boldsymbol{N},$$

pero la sucesión X no es ni monótonamente decreciente ni creciente. (De hecho, los términos con índice impar forman una sucesión creciente y aquellos con índice par forman una sucesión decreciente.) Dado que los términos de la sucesión se forman promediando, es fácil ver que

$$|x_n - x_{n+1}| = \frac{1}{2^{n-1}} \quad \text{para} \quad n \in \boldsymbol{N}.$$

De modo que si $m > n$, se usa la desigualdad del triángulo para obtener

$$|x_n - x_m| \leq |x_n - x_{n+1}| + \cdots + |x_{m-1} - x_m|$$
$$= \frac{1}{2^{n-1}} + \cdots + \frac{1}{2^{m-2}} = \frac{1}{2^{n-1}}\left(1 + \frac{1}{2} + \cdots + \frac{1}{2^{m-n-1}}\right) < \frac{1}{2^{n-2}}.$$

Dada $\varepsilon > 0$, si n se escoge de tal magnitud que $1/2^n < \varepsilon/4$ y si $m \geq n$, se deduce que

$$|x_n - x_m| < \varepsilon.$$

Por lo tanto, X es una sucesión de Cauchy en $\boldsymbol{R}$ y, por el criterio de Cauchy la sucesión X converge a un número x. Para calcular el límite se advierte que tomando el límite según la definición se obtiene un resultado válido pero poco informativo,

$$x = \tfrac{1}{2}(x + x).$$

Sin embargo, dado que converge la sucesión X, también converge la subsucesión con índices impares. Por medio de inducción se puede establecer que

$$x_1 = 1, \qquad x_3 = 1 + \frac{1}{2}, \qquad x_5 = 1 + \frac{1}{2} + \frac{1}{2^3}, \ldots,$$
$$x_{2n+1} = 1 + \frac{1}{2} + \frac{1}{2^3} + \cdots + \frac{1}{2^{2n-1}}, \ldots$$

Se deduce que

$$x_{2n+1} = 1 + \frac{1}{2}\left(1 + \frac{1}{4} + \cdots + \frac{1}{4^{n-1}}\right)$$
$$= 1 + \frac{1}{2} \cdot \frac{1 - 1/4^n}{1 - 1/4} = 1 + \frac{2}{3}\left(1 - \frac{1}{4^n}\right).$$

Por lo tanto, la subsucesión con índices impares converge a $\frac{5}{3}$; por lo que toda la sucesión tiene el mismo límite.

(b) Sea $X = (x_n)$ una sucesión real dada por

$$x_1 = \frac{1}{1!}, \quad x_2 = \frac{1}{1!} - \frac{1}{2!}, \ldots, \quad x_n = \frac{1}{1!} - \frac{1}{2!} + \cdots + \frac{(-1)^{n+1}}{n!}, \ldots$$

Como esta sucesión no es monótona, no es posible aplicar directamente el teorema de convergencia monótona. Observe que si $m > n$, entonces

$$x_m - x_n = \frac{(-1)^{n+2}}{(n+1)!} + \frac{(-1)^{n+3}}{(n+2)!} + \cdots + \frac{(-1)^{m+1}}{m!}.$$

Recordando que $2^{r-1} \le r!$, se obtiene

$$|x_m - x_n| \le \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots + \frac{1}{m!}$$
$$\le \frac{1}{2^n} + \frac{1}{2^{n+1}} + \cdots + \frac{1}{2^{m-1}} < \frac{1}{2^{n-1}}.$$

Por lo tanto, la sucesión es una sucesión de Cauchy en $\boldsymbol{R}$.

(c) Si $X = (x_n)$ es una sucesión en R definida por

$$x_n = \frac{1}{1} + \frac{1}{2} + \cdots + \frac{1}{n} \qquad \text{para} \quad n \in \mathbf{N},$$

y si $m > n$, entonces

$$x_m - x_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{m}.$$

Dado que cada uno de estos $m - n$ términos excede a $1/m$, dicha diferencia excede a $(m-n)/m = 1 - n/m$. En particular si $m = 2n$, se tiene

$$x_{2n} - x_n > \tfrac{1}{2}$$

Esto prueba que X no es una sucesión de Cauchy; por lo tanto, se concluye que X es divergente. (Se ha demostrado que la "serie armónica" es divergente.)

Ejercicios

16.A. Sea $x_1 \in \boldsymbol{R}$ tal que satisfaga $x_1 > 1$ y sea $x_{n+1} = 2 - 1/x_n$ para $n \in \boldsymbol{N}$. Probar que la sucesión (x_n) es monótona y acotada. ¿Cuál es su límite?

16.B. Sean $y_1 = 1$ y $y_{n+1} = (2 + y_n)^{1/2}$ para $n \in \boldsymbol{N}$. Demostrar que es monótona y acotada. ¿Cuál es su límite?

16.C. Sean $a > 0$ y $z_1 > 0$. Defínase $z_{n+1} = (a + z_n)^{1/2}$ para $n \in \boldsymbol{N}$. Probar que (z_n) converge.

16.D. Si a satisface $0 < a < 1$, demostrar que la sucesión $X = (a^n)$ es convergente. Dado que $Y = (a^{2n})$ es una subsucesión, se tiene $\lim X = \lim Y = (\lim X)^2$, y que $\lim X = 0$.

16.E. Demostrar que toda sucesión en $\boldsymbol{R}$ tiene una subsucesión monótona creciente o bien una sucesión monótona decreciente.

16.F. Usar el ejercicio 16.E para demostrar el teorema de Bolzano-Weierstrass para sucesiones en $\boldsymbol{R}$.

16.G. Determinar la convergencia o divergencia de la sucesión (x_n), en dónde

$$x_n = \frac{1}{n+1} + \frac{1}{n+2} + \cdots + \frac{1}{2n} \qquad \text{para} \qquad n \in \boldsymbol{N}.$$

16.H. Sean $X = (x_n)$ y $Y = (y_n)$ sucesiones en $\boldsymbol{R}^p$ y sea $Z = (z_n)$ la sucesión "intercalar" definida por $z_1 = x_1$, $z_2 = y_1, \ldots, z_{2n} = x_n$, $z_{2n+1} = y_n, \ldots$ s verdad que Z es convergente si y sólo si X y Y son convergentes y $\lim X = \lim Y$?

16.I. Demostrar directamente que las siguientes son sucesiones de Cauchy:

(a) $\left(\frac{1}{n}\right)$, (b) $\left(\frac{n+1}{n}\right)$, (c) $\left(1 + \frac{1}{1!} + \cdots + \frac{1}{n!}\right)$.

16.J. Demostrar directamente que las siguientes no son sucesiones de Cauchy:

(a) $((-1)^n)$, (b) $(n + (-1)^n/n)$, (c) (n^2).

16.K. Sea $X = (x_n)$ una sucesión de números reales estrictamente positivos sea $\lim (x_{n+1}/x_n) = L$, ea $0 < \varepsilon < L$. Demostrar que existen $A > 0$, $B > 0$, $K \in \boldsymbol{N}$ es que $A(L-\varepsilon)^n \leq x_n \leq B(L+\varepsilon)^n$ para $n \geq K$. Después demostrar que es $\lim (x_n^{1/n}) = L$.

16.L. Aplicar el ejemplo 16.3*(d)* y el ejercicio anterior a la sucesión $(n^n/n!)$ para demostrar que $\lim (n/(n!)^{1/n}) = e$.

16.M. Establecer la convergencia y los límites de las siguientes sucesiones:

(a) $((1+1/n)^{n+1})$, (b) $((1+1/2n)^n)$,
(c) $((1+2/n)^n)$, (d) $((1+1/(n+1))^{3n})$.

16.N. Sea $0 < a_1 < b_1$ y defínase, para $n \in \boldsymbol{N}$,

$$a_{n+1} = (a_n b_n)^{1/2}, \qquad b_{n+1} = \tfrac{1}{2}(a_n + b_n).$$

Por inducción demostrar que $a_n < b_n$. Demostrar que (a_n) y (b_n) convergen al mismo límite.

16.0. Dar una demostración para el teorema de intersección de Cantor 11-4 tomando un punto $x_n \in F_n$ y aplicando el teorema de Bolzano-Weierstrass 16.4.

16.P. Dar una demostración para el teorema del punto nidificado 11.6 usando el teorema de Bolzano-Weierstrass 16.4.

16.Q. Demostrar que si K_1 y K_2 son subconjuntos compactos de $\boldsymbol{R}^p$ entonces existen puntos $x_1 \in K_1$, $x_2 \in K_2$ tales que si $z_1 \in K_1$, $z_2 \in K_2$, entonces $\|z_1 - z_2\| \geq \|x_1 - x_2\|$.

Proyecto

16.α. En este proyecto, sean m, c y c_0 las letras que designan a las colecciones de sucesiones reales iniroducidas en el proyecto 15.β y sea d la métrica definida en la parte *(d)* de dicho proyecto.

(a) Si $r \in \boldsymbol{I}$ y $r = 0\,.\,r_1 r_2 \cdots r_n \cdots$ es su expansión decimal, considérese al elemento $X_r = (r_n)$ en m. Deducir que hay un subconjunto no contable A de m tal que si X_r y X_s son elementos distintos de A entonces $d(X_r, X_s) \geq 1$.

(b) Suponga que B es un subconjunto de c con la propiedad de que si X y Y son distintos elementos de B entonces $d(X, Y) \geq 1$. Demostrar que B es un conjunto contable.

(c) Si $j \in \boldsymbol{N}$, sea $Z_j = (z_{nj} : n \in \boldsymbol{N})$ la sucesión cuyos primeros j elementos son 1 y cuyos elementos restantes son 0. Observe que Z_j pertenece a cada uno de los espacios métricos m, c y c_0 y que $d(Z_j, Z_k) = 1$ para $j \neq k$. Demostrar que la sucesión $(Z_j : j \in \boldsymbol{N})$ es monótona en el sentido de que cada sucesión coordenada $(z_{nj} : j \in \boldsymbol{N})$ es monótona. Demostrar que la sucesión (Z_j) no converge con respecto a la métrica d en ninguno de los tres espacios.

(d) Probar que hay una sucesión (X_j) en m, c y c_0 que es acotada (en el sentido de que existe una constante K tal que $d(X_j, 0) \leq K$ para toda $j \in \boldsymbol{N}$) pero que no posee ninguna subsucesión convergente.

(e) (Si d es una métrica en un conjunto M, se dice que una sucesión (X_j) en M es una **sucesión de Cauchy** si para toda $\varepsilon > 0$ existe $K(\varepsilon) \in \boldsymbol{N}$ tal que $d(X_j, X_k) < \varepsilon$ siempre que $j, k \geq K(\varepsilon)$. Se dice que M es **completo** con respecto a d cuando toda sucesión de Cauchy en M converge a un elemento de M.) Demostrar que los conjuntos m, c y c_0 son completos con respecto a la métrica d que se ha estado considerando.

(f) Sea f la colección de todas las sucesiones reles que tienen solamente un número finito de elementos distintos de cero y defínase a d igual que se hizo antes. Probar que d es una métrica en f pero que f no es completo con respecto a d.

Sección 17 Sucesiones de funciones

En las tres secciones anteriores se consideró la convergencia de sucesiones de elementos de $\boldsymbol{R}^p$; en esta sección se considerarán sucesiones de *funciones*. Después de dar algunos datos preliminares sencillos se introducirá el concepto un tanto sutil pero básico de convergencia uniforme de una sucesión de funciones.

Sea $D \subseteq \boldsymbol{R}^p$ dado y suponga que para cada número natural $n \in \boldsymbol{N}$ hay una función f_n con dominio D y rango en $\boldsymbol{R}^q$; se dirá que (f_n) es una *sucesión de funciones* en $D \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^q$. Se debe comprender que para cualquier punto x en D dicha sucesión de funciones da una sucesión de elementos en $\boldsymbol{R}^q$; específicamente, la sucesión.

(17.1) $$(f_n(x))$$

que se obtiene calculando cada una de las funciones en x. Para ciertos puntos x en D, la sucesión (17.1) puede converger y para otros puntos x en D la sucesión puede divergir. Para cada uno de los puntos x para los cuales la sucesión (17.1) converge, por el teorema 14.5, existe un punto de $\boldsymbol{R}^q$ determinado de manera única. En general, el valor de este límite, cuando exista, dependerá de la elección del punto x. De esta manera, surge una función cuyo dominio consta de todos los puntos x en $D \subseteq \boldsymbol{R}^p$ para los cuales la sucesión (17.1) converge en $\boldsymbol{R}^q$.

Con estas palabras introductorias se resumirá una definición formal de convergencia de una sucesión de funciones.

17.1 DEFINICION. Sea (f_n) una sucesión de funciones en $D \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^q$, sea D_0 un subconjunto de D y sea f una función con un dominio que contiene a D_0 y rango en $\boldsymbol{R}^q$. Se dice que la **sucesión (f_n) converge en D_0 a f** si, para cada x en D_0 la sucesión $(f_n(x))$ converge $\boldsymbol{R}^q$ a $f(x)$. En este caso, a la función f se la llama **el límite en D_0 de la sucesión (f_n)**. Cuando existe dicha función f se dice que la sucesión (f_n) **converge a f en D_0** o simplemente que la sucesión es **convergente en D_0**.

Del teorema 14.5 se deduce que, excepto por algún posible cambio en el dominio D_0 la función límite se determina de manera única. Por lo común se elige D_0 como el máximo conjunto posible, es decir el conjunto de todas las x en D para las cuales (17.1) converge. Para expresar simbólicamente que la sucesión (f_n) converge en D_0 a f algunas veces se escribe

$$f = \lim (f_n) \quad \text{on } D_0, \qquad \text{ó} \qquad f_n \to f \quad \text{on } D_0.$$

En seguida se verán algunos ejemplos acerca de esta idea. Para simplificar se tratará el caso especial $p = q = 1$.

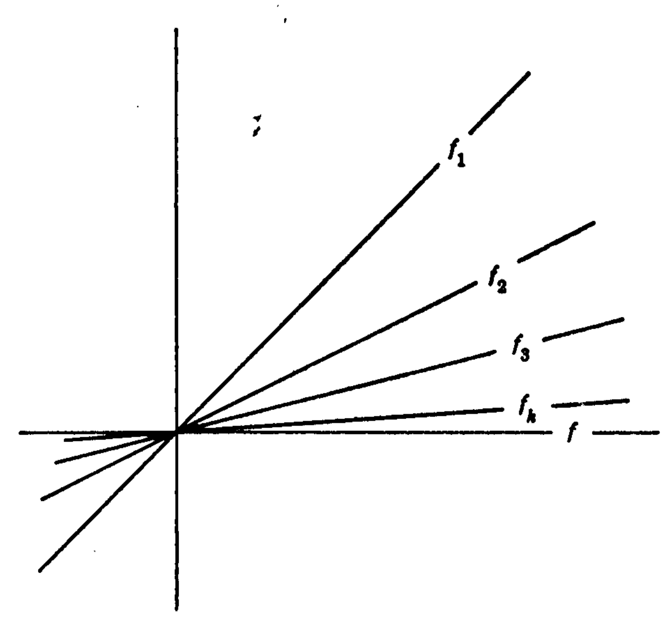


Figura 17.1

17.2 EJEMPLOS. *(a)* Para cada número natural n, defínase f_n para x en $D = \boldsymbol{R}$ como $f_n(x) = x/n$. Defínase f, para toda x en $D = \boldsymbol{R}$ como $f(x) = 0$. (figura 17.1) La afirmación de que la sucesión (f_n) converge en R a f es equivalente a la afirmación de que para cada número real x la sucesión numérica (x/n) converge a O. Para ver que este es el caso se aplican el ejemplo 14.8*(a)* y el teorema 15.6*(b)*.

(b) Sea $D = \{x \in \boldsymbol{R} : 0 \le x \le 1\}$ y para cada número natural n, defínase f_n como $f_n(x) = x^n$ para toda x en D defínase a f como

$$f(x) = 0, \qquad 0 \le x < 1,$$
$$= 1, \qquad x = 1.$$

(figura 17.2.) Es claro que cuando $x = 1$, entonces $f_n(x) = f_n(1) = 1^n = 1$ de manera que $f_n(1) \to f(1)$. En el ejemplo 14.8*(c)* se probó que si $0 \le x < 1$, entonces $f_n(x) = x^n \to 0$. Por lo tanto, se concluye que (f_n) converge en D a f. (No es difícil demostrar que si $x > 1$ entonces $(f_n(x))$ no converge en lo absoluto.)

(c) Sea $D = \boldsymbol{R}$ y para cada número natural n, sea f_n la función definida para x en D por medio de

$$f_n(x) = \frac{x^2 + nx}{n},$$

y sea $f(x) = x$ (figura 17.3.) Dado que $f_n(x) = (x^2/n) + x$, del ejemplo 14.8*(a)* y del teorema 15.6*(b)* se deduce que $(f_n(x))$ converge a $f(x)$ para toda $x \in \boldsymbol{R}$.

(d) Sea $D = \boldsymbol{R}$ y, para cada número natural n, defínase f_n como $f_n(x) = (1/n)$ sen $(nx + n)$ (figura 17.4.) (En este caso no es necesario una definición rigurosa de la función seno; de hecho, todo lo que se necesita es que $|\text{sen}\, y| \le 1$ para cualquier número real y.) Si f se define como la función

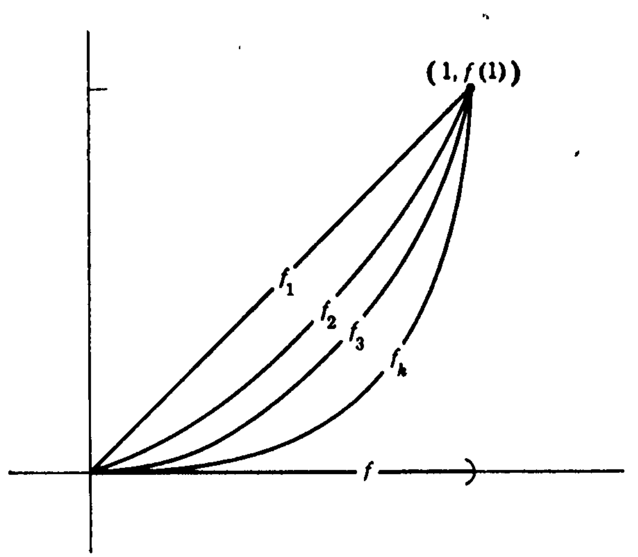


Figura 17.2

cero $f(x)=0$, $x \in \boldsymbol{R}$, entonces $f = \lim(f_n)$ En efecto, para cualquier número real x se tiene

$$|f_n(x) - f(x)| = \frac{1}{n}\,|\text{sen}\,(nx+n)| \leq \frac{1}{n}.$$

Si $\varepsilon > 0$, existe un número natural $K(\varepsilon)$ tal que si $n \geq K(\varepsilon)$, entonces $1/n < \varepsilon$. Por lo que para dicha n se concluye que

$$|f_n(x) - f(x)| < \varepsilon$$

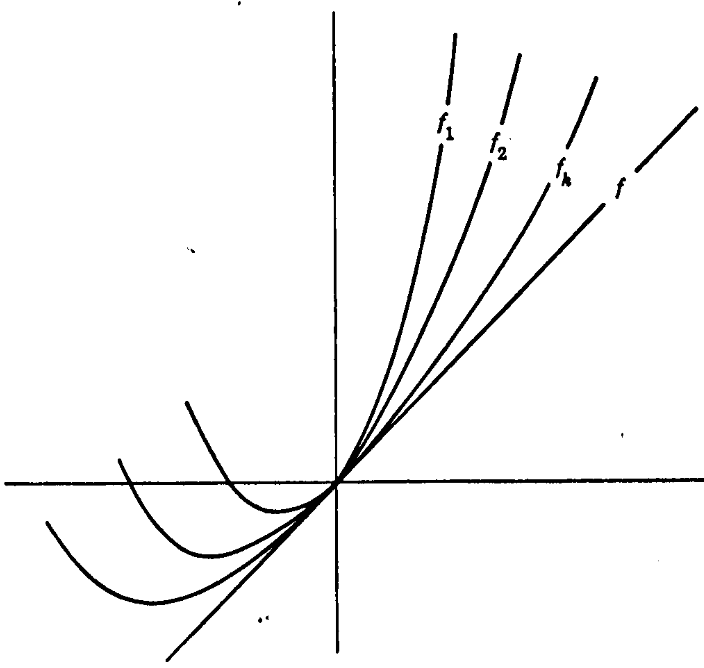


Figura 17.3

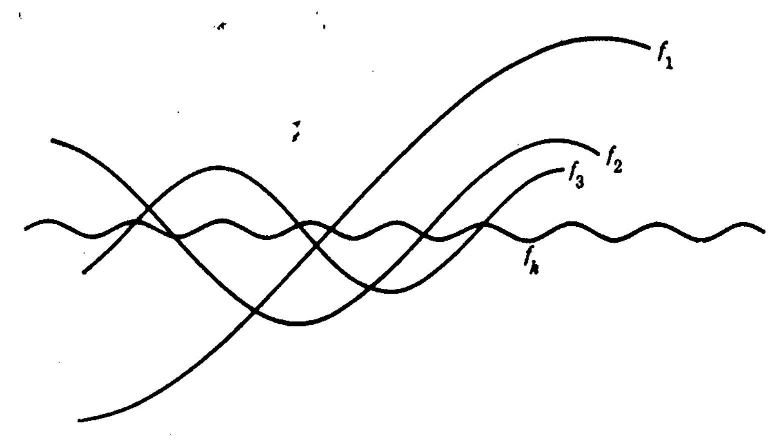


Figura 17.4

para cualquier valor de x. Por lo tanto, se deduce que la sucesión (f_n) converge a f. (Obsérvese que eligiendo n lo suficientemente grande se pueden hacer las diferencias $|f_n(x)-f(x)|$ menores que ε para todo valor de x simultáneamente).

La siguiente reafirmación de la definición 17.1 se formula en parte para reforzar la definición 17.1 y en parte para prepararle el camino al concepto tan importante de convergencia uniforme.

17.3 LEMA. *Una sucesión (f_n) de funciones en $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ converge a una función f en $D_0 \subseteq D$ si y sólo si para cada $\varepsilon > 0$ y cada x en D_0 hay un número natural $K(\varepsilon, x)$ tal que para toda $n \geq K(\varepsilon, x)$,*

$$\|f_n(x)-f(x)\| < \varepsilon. \tag{17.2}$$

Dado que esto es simplemente una reformulación de la definición 17.1, no se repetirán los detalles de la demostración, pero se dejan al lector como ejercicio. Sólo se hace la aclaración de que el valor de n que se requiere en la desigualdad (17.2) dependerá, en general, de $\varepsilon > 0$ *así como* de $x \in D_0$. El lector cuidadoso ya se habrá dado cuenta de que en los ejemplos 17.2*(a-c)* el valor de n que se requiere para obtener (17.2) depende de ambos: $\varepsilon > 0$ y $x \in D_0$. Sin embargo, en el ejemplo 17.2(*d*) la desigualdad (17.2) se puede cumplir para toda x en D_0 siempre que n se elija lo suficientemente grande pero dependiendo tan sólo de ε .

Es precisamente esta diferencia, un tanto sutil, la que permite distinguir entre los conceptos de convergencia "ordinaria" de una sucesión de funciones (en el sentido de la definición 17.1) y Convergencia "uniforme" que en seguida se define.

17.4 DEFINICION. Una sucesión (f_n) de funciones en $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ **converge uniformemente** en un subconjunto D_0 de D a una función f cuando para cada $\varepsilon > 0$ hay un número natural $K(\varepsilon)$ (dependiente de E pero **no** de $x \in D_0$) tal que para toda $n \geq K(\varepsilon)$ y $x \in D_0$,

(17.3) $$\|f_n(x) - f(x)\| < \varepsilon.$$

En este caso se dice que la sucesión es **uniformemente convergente** en D_0 (figura 17.5.)

Se deduce en forma directa que si la sucesión (f_n) es uniformemente convergente en D_0 a f entonces esta sucesión de funciones también converge a f en el sentido de la definición 17.1. Se puede ver que lo inverso no es válido mediante un análisis cuidadoso de los ejemplos 17.2 *(a-c)*. Se darán otros ejemplos más adelante. Antes de continuar es conveniente establecer una condición necesaria y suficiente para que la sucesión (f_n) **deje** de convergir uniformemente en D_0 a f.

17.5 LEMA. *Una sucesión (f_n) no converge uniformemente en D_0 a f si y sólo si para alguna $\varepsilon_0 > 0$ hay una subsucesión (f_{n_k}) de (f_n) y una sucesión (x_k) en D_0 tales que*

(17.4) $$\|f_{n_k}(x_k) - f(x_k)\| \geq \varepsilon_0 \qquad \textit{for} \quad k \in \mathbf{N}.$$

Para demostrar este resultado se requiere simplemente que el lector anule la definición 17.4. Se le deja al lector como ejercicio indispensable. El lema anterior es provechoso para probar que los ejemplos 17.2 (*a-c*) no convergen uniformemente en los conjuntos dados D_0.

17.6 EJEMPLOS. *(a)* Considere el ejemplo 17.2(*a*). Si $n_k = k$ y $x_k = k$, entonces $f_k(x_k) = 1$ de tal manera que

$$|f_k(x_k) - f(x_k)| = |1 - 0| = 1.$$

Esto prueba que la sucesión (f_n) no converge uniformemente en $\boldsymbol{R}$ a f.

(b) Considere el ejemplo 17.2(*b*). Si $n_k = k$ y $x_k = (\frac{1}{2})^{1/k}$, entonces

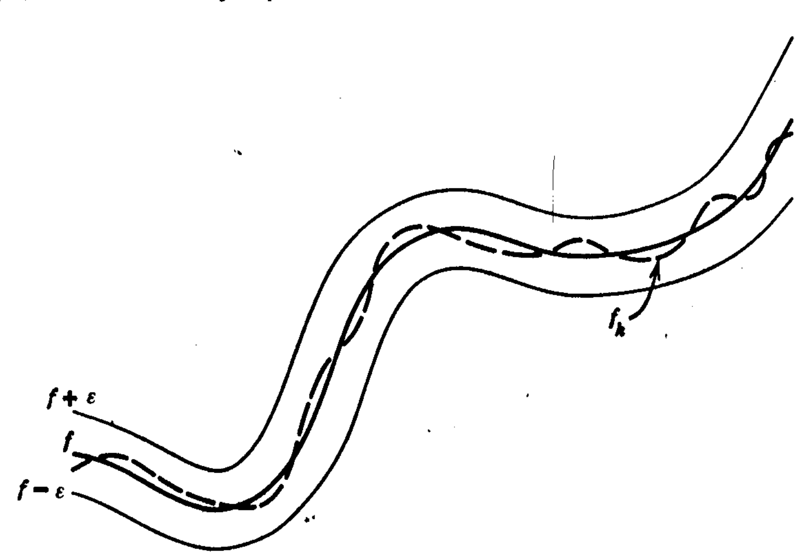


Figura 17.5

$$|f_k(x_k) - f(x_k)| = |f_k(x_k)| = \tfrac{1}{2}.$$

Por lo tanto, se deduce que la sucesión (f_n) no converge uniformemente en $[0, 1]$ a f.

(c) Considere el ejemplo 17.2(*c*). Si $n_k = k$ y $x_k = k$, entonces

$$|f_k(x_k) - f(x_k)| = k,$$

probando que (f_k) no converge uniformemente in $\boldsymbol{R}$ a f.

(d) Considere el ejemplo 17.2*(d)*. Dado que

$$|f_n(x) - f(x)| \leq 1/n$$

para toda x en $\boldsymbol{R}$, la sucesión (f_n) converge uniformemente en $\boldsymbol{R}$ a f.

La norma uniforme

Al analizar la convergencia uniforme con frecuencia es conveniente emplear cierta norma en un espacio vectorial de funciones.

Si $D \subseteq \boldsymbol{R}^p$ y $f: D \to \boldsymbol{R}^q$, se dice que f es **acotada** siempre que exista $M > 0$ tal que $\|f(x)\| \leq M$ para toda $x \in D$. Si deduce $f: D \to \boldsymbol{R}^q$ es acotada, entonces se deduce que el número $\|f\|_D$ definido por medio de

$$\|f\|_D = \sup \{\|f(x)\| : x \in D\} \tag{17.5}$$

existe en $\boldsymbol{R}$. (Se puede ver que la norma del lado derecho de esta ecuación es la norma en el espacio $\boldsymbol{R}^q$.).

17.7 DEFINICION. Si $D \subseteq \boldsymbol{R}^p$, entonces la colección de todas las funciones acotadas en D a $\boldsymbol{R}^q$ se designa por medio de $B_{pq}(D)$ o (cuando p y q se sobreentienden) por medio de $B(D)$.

En el espacio $B_{pq}(D)$ se define la **suma vectorial** de dos funciones f, g y la **multiplicación escalar** de $c \in R$ y f como

$$(f + g)(x) = f(x) + g(x), \qquad (cf)(x) = cf(x) \tag{17.6}$$

para toda $x \in D$. La función cero se define como la función $0: D \to \boldsymbol{R}^q$ definida para toda $x \in D$ por medio de $0(x) = 0$. En seguida se relaciona esta terminología con los conceptos que se dieron en la sección 8

17.8 LEMA. *(a) El conjunto $B_{pq}(D)$ es un espacio vectorial bajo las operaciones vectoriales definidas en la ecuación* (17.6).

(b) La función $f \mapsto \|f\|_D$ definida en $B_{pq}(D)$ en la ecuación (17.5) *es una norma en $B_{pq}(D)$.*

DEMOSTRACION. La demostración de (*a*) sólo requiere cálculos rutinarios.

Para demostrar (*b*), es necesario establecer las cuatro propiedades de una norma que se dan en la definición 8.5 (*i*) De (17.5) es claro que $\|f\|_D \geq 0$. *(ii)* Claramente, $\|0\|_D = \sup \{\|0(x)\| : x \in D\} = 0$ A la inversa, si $\|f\|_D = 0$, enton-

ces como $0 \le \|f(x)\| \le \|f\|_D = 0$, se deduce que $\|f(x)\| = 0$ y por lo tanto $f(x) = 0$ para toda $x \in D$ de manera que $f = 0$. *(iii)* El hecho de que $\|cf\|_D = |c|\, \|f\|_D$ se puede ver fácilmente *(iv)* Dado que

$$\begin{aligned}\|(f+g)(x)\| = \|f(x)+g(x)\| &\le \|f(x)\| + \|g(x)\| \\ &\le \|f\|_D + \|g\|_D\end{aligned}$$

para toda $x \in D$, se deduce que $\|f\|_D + \|g\|_D$ es una cota superior para el conjunto $\{\|(f+g)(x)\| : x \in D\}$. Por lo tanto, se tiene

$$\begin{aligned}\|f+g\|_D &= \sup\{\|(f+g)(x)\| : x \in D\} \\ &\le \|f\|_D + \|g\|_D.\end{aligned}$$

Q.E.D.

Algunas veces a la norma $f \mapsto \|f\|_D$ se le llama **norma uniforme** (o **norma suprema**) **en** $B_{pq}(D)$. En seguida se demostrará que la convergencia uniforme de funciones en $B_{pq}(D)$ es equivalente a la convergencia en la norma uniforme.

17.9 TEOREMA. *Una sucesión* (f_n) *en* $B_{pq}(D)$ *converge uniformemente en* D *a* $f \in B_{pq}(D)$ *si y sólo si*

$$\|f_n - f\|_D \to 0.$$

DEMOSTRACION. Si la sucesión (f_n) converge uniformemente a f en D, entonces para cualquier $\varepsilon > 0$ hay un número natural $K(\varepsilon)$ tal que si $n \ge K(\varepsilon)$ y $x \in D$, $\|f_n(x) - f(x)\| < \varepsilon$. Esto implica que

$$\|f_n - f\|_D = \sup\{\|(f_n - f)(x)\| : x \in D\} \le \varepsilon.$$

Dado que $\varepsilon > 0$ es arbitraria, se tiene $\|f_n - f\|_D \to 0$.

Inversamente, si $\|f_n - f\|_D \to 0$, entonces dada $\varepsilon > 0$ existe $K(\varepsilon)$ tal que si $n \ge K(\varepsilon)$ $\|f_n - f\|_D \le \varepsilon$. Esto implica que si $x \in D$, entonces

$$\|f_n(x) - f(x)\| = \|(f_n - f)(x)\| \le \|f_n - f\|_D \le \varepsilon.$$

Por lo tanto, la sucesión (f_n) converge uniformemente en D a f. Q.E.D.

Se va a explicar el uso de este lema como herramienta para analizar una sucesión de funciones de convergencia uniforme. Primero se observa que la norma sólo se ha definido para funciones acotadas; por lo tanto, se puede usar (al menos directamente) sólo cuando la sucesión consta de funciones acotadas

17.10 EJEMPLOS. *(a)* El lema 17.9 no se puede aplicar al ejemplo tomado en 17.2(*a*) y 17.6(*a*) ya que las funciones f_n, definidas como $f_n(x) = x/n$, no están acotadas en $\boldsymbol{R}$, que se dio como dominio. Para mayor claridad, se cambiará el dominio para obtener una sucesión acotada en el nuevo dominio. Por conveniencia tómese $E = [0, 1]$. A pesar de que en el dominio $\boldsymbol{R}$ la sucesión (x/n) no convergía uniformemente a la función cero

(como se vio en el ejemplo 17.6(*a*)), la convergencia es uniforme en $E=[0,1]$. Para ver esto se calcula

$$\|f_n-f\|_E=\sup\left\{\left|\frac{x}{n}-0\right|:0\leq x\leq 1\right\}=\frac{1}{n};$$

por lo tanto $\|f_n-f\|_E=1/n\to 0$.

(*b*) Ahora se considera la sucesión que se vio en los ejemplos 17.2(*b*) y 17.6(*b*). En este caso $D=[0,1]$, $f_n(x)=x^n$, y la función límite f es igual a 0 para $0\leq x<1$ e igual a 1 para x = 1. Calculando la norma de la diferencia f_n-f, se obtiene

$$\|f_n-f\|_D=\sup\begin{cases}x^n, & 0\leq x<1\\ 0, & x=1\end{cases}=1 \qquad \text{para} \quad n\in\mathbf{N}.$$

Dado que esta norma no converge a cero, se deduce que la sucesión (f_n) no converge uniformemente en $D=[0,1]$ to f. Esto confirma las consideraciones anteriores.

(*c*) Considere el ejemplo 17.2(*c*). Una vez más, no se puede aplicar el lema 17.9, ya que las funciones no son acotadas. De nuevo se elige un dominio más pequeño, sea $E=[0,a]$ con $a>0$. Dado que

$$|f_n(x)-f(x)|=\left|\frac{x^2+nx}{n}-x\right|=\frac{x^2}{n},$$

se tiene

$$\|f_n-f\|_E=\sup\{|f_n(x)-f(x)|:0\leq x\leq a\}=\frac{a^2}{n}.$$

De donde, la sucesión converge uniformemente a f en el intervalo $[0,a]$. (¿por qué no contradice esto al resultado obtenido en el ejemplo 17.6(*c*)?)

(*d*) Haciendo referencia al ejemplo 17.2(*d*), considérese la función $f_n(x)=(1/n)$ sen $(nx+n)$ en $D=\mathbf{R}$. Aquí, la función límite $f(x)=0$ para toda $x\in D$. Para establecer la convergencia uniforme de esta sucesión, observe que

$$\|f_n-f\|_D=\sup\{(1/n)\,|\text{sen}\,(nx+n)|:x\in\mathbf{R}\}$$

Pero como $|\text{sen}\ y|\leq 1$, se concluye que $\|f_n-f\|_D=1/n$. Por lo tanto, (f_n) converge uniformemente en R, como se había asegurado en el ejemplo 17.6(*d*).

Uno de los aspectos más útiles de la norma es que facilita la formulación de un criterio de Cauchy para la convergencia uniforme de una sucesión de funciones acotadas.

17.11 CRITERIO DE CAUCHY PARA CONVERGENCIA UNIFORME. *Sea (f_n) una sucesión de funciones en $\mathbf{B}_{pq}(D)$. Entonces, hay una función $f\in\mathbf{B}_{pq}(D)$ a la cual (f_n) converge uniformemente en D si y sólo si para cada $\varepsilon>0$ hay un número natural $M(\varepsilon)$ tal que para todas $m,n\geq M(\varepsilon)$, $\|f_m-f_n\|_D<\varepsilon$.*

DEMOSTRACION. Suponga que la sucesión (f_n) converge uniformemente en D a una función $f \in B_{pq}(D)$. Entonces, para $\varepsilon > 0$ hay un número natural $K(\varepsilon)$ tal que si $n \geq K(\varepsilon), \|f_n - f\|_D < \varepsilon/2$. Por lo que si, $m, n \geq K(\varepsilon)$, se concluye que

$$\|f_m - f_n\|_D \leq \|f_m - f\|_D + \|f - f_n\|_D < \varepsilon.$$

Inversamente, suponga que se satisface el criterio de Cauchy y que para $\varepsilon > 0$ hay un número natural $M(\varepsilon)$ al que $\|f_m - f_n\|_D < \varepsilon$ cuando $m, n \geq M(\varepsilon)$. Ahora, para cada $x \in D$ se tiene

(17.6) $$\|f_m(x) - f_n(x)\| \leq \|f_m - f_n\|_D < \varepsilon \qquad \text{para} \quad m, n \geq M(\varepsilon).$$

Por lo que la sucesión $(f_n(x))$ es una sucesión de Cauchy en $\boldsymbol{R}^q$ y converge a algún elemento de $\boldsymbol{R}^q$. Defínase f para x en D como

$$f(x) = \lim (f_n(x)).$$

A partir de (17.6) se concluye que si m es un número natural fijo que satisface $m \geq M(\varepsilon)$ y si n es cualquier número natural con $n \geq M(\varepsilon)$, entonces para toda x en D se tiene

$$\|f_m(x) - f_n(x)\| < \varepsilon.$$

Si se aplica el lema 15.8 se deduce que si $m \geq M(\varepsilon)$ y $x \in D$, entonces $\|f_m(x) - f(x)\| \leq \varepsilon$. Dado que f_m es una función acotada fácilmente se deduce de aquí (¿cómo?) que f es acotada y por lo tanto pertenece a $B_{pq}(D)$. Más aún, se concluye que (f_n) converge uniformemente a f en D. Q.E.D.

Ejercicios

En estos ejercicios se puede hacer uso de las propiedades elementales de funciones trigonométricas y exponenciales estudiadas en cursos anteriores.

17.A. Para cada $n \in \boldsymbol{N}$, defínase f_n para $x > 0$ como $f_n(x) = 1/(nx)$. ¿Para qué valores de x existe $\lim (f_n(x))$?

17.B. Para cada $n \in \boldsymbol{N}$, defínase g_n para $x \geq 0$ por medio de la fórmula

$$\begin{aligned} g_n(x) &= nx, & 0 \leq x \leq 1/n, \\ &= \frac{1}{nx}, & 1/n < x, \end{aligned}$$

Demostrar que $\lim (g_n(x)) = 0$ para toda $x > 0$.

17.C. Demostrar que $\lim ((\cos \pi x)^{2n})$ existe para todos los valores de x. ¿Cuál es el límite?

17.D. Demostrar que si f_n se define en $\boldsymbol{R}$ por medio de

$$f_n(x) = \frac{nx}{1 + n^2x^2},$$

entonces, (f_n) converge en $\boldsymbol{R}$.

17.E. Defínase h_n en el intervalo $\boldsymbol{I} = [0, 1]$ por medio de la fórmula

i146 (1478x2460x2 pcx)

$$h_n(x) = 1 - nx, \qquad 0 \le x \le 1/n,$$
$$= 0, \qquad 1/n < x \le 1.$$

Demostrar que lim (h_n) existe en I.

17.F. Defínase (g_n) en I por medio de

$$g_n(x) = nx, \qquad 0 \le x \le 1/n,$$
$$= \frac{n}{n-1}(1-x), \qquad 1/n < x \le 1.$$

Demostrar que lim g_n existe en I.

17.G. Demostrar que si f_n se define en R por medio de

$$f_n(x) = \frac{2}{\pi} \text{ Arc tan } (nx),$$

entonces f = lim (f_n) existe en R. De hecho, el límite está dado por

$$f(x) = 1, \qquad x > 0,$$
$$= 0, \qquad x = 0,$$
$$= -1, \qquad x < 0.$$

17.H. Demostrar que lim (e^{-nx}) existe para $x \ge 0$ también la existencia de lim (xe^{-nx}).

17.I. Suponga que (x_n) es una sucesión convergente de puntos que se encuentra, igual que su límite x, en el conjunto $D \subseteq R^p$. Suponga que (f_n) converge en D a la función f. ¿Es verdad que $f(x) = \lim (f_n(x_n))$?

17.J. Considere el ejercicio anterior con la hipótesis adicional de que la convergencia de (f_n) es uniforme en D.

17.K. Demostrar que la convergencia en el ejercicio 17.A. no es uniforme en todo el conjunto de convergencia pero es uniforme para $x \ge 1$.

17.L. Demostrar que la convergencia en el ejercicio 17.B no es uniforme en el dominio $x \ge 0$, pero es uniforme en un conjunto $x \ge c$, en donde $c > 0$.

17.M. ¿Es uniforme en R la convergencia del ejercicio 17.D?

17.N. ¿Es uniforme en I la convergencia del ejercicio 17.E?

17.O. ¿Es uniforme en I la convergencia del ejercicio 17.F? ¿es uniforme en $[c, 1]$ para $c > 0$?

17.P. ¿Converge uniformemente para (xe^{-nx}) la sucesión $x \ge 0$?

17.Q. ¿Converge uniformemente para (x^2e^{-nx}) la sucesión $x \ge 0$?

17.R. Sea (f_n) una sucesión de funciones que converge en D a una función f. Si A y B son subconjuntos de D y se sabe que la convergencia es uniforme en A y también en B, demostrar que la convergencia es uniforme en $A \cup B$.

17.5. Dar un ejemplo de una sucesión (f_n) en $B_{pq}(I)$ tal que $\|f_n\|_I \le 1$ para toda $n \in N$ que no tenga alguna subsucesión uniformemente convergente. (De donde se explica que el teorema de Bolzano-Weierstrass no se cumpla en $B_{pq}(I)$.)

Sección 18 El límite superior

En la sección 6 se introdujo el concepto del supremo de un conjunto acotado no vacío de números reales que se ha utilizado varias veces de manera



Figura 18.1

importante. Sin embargo, tratándose de un conjunto infinito acotado $S \subseteq \boldsymbol{R}$ algunas veces también resulta interesante considerar el mayor punto de acumulación s^* de S. Este punto s^* es el ínfimo de todos los números reales que son excedidos a lo más por un número finito de elementos de S. Se adaptará este concepto a sucesiones acotadas en $\boldsymbol{R}$ para obtener el concepto, a menudo útil de "límite superior".

18.1 DEFINICION. Sea $X = (x_n)$ una sucesión acotada en $\boldsymbol{R}$.

(a) El **límite superior** de X, que se denota mediante el símbolo

$$\limsup X, \qquad \limsup (x_n), \qquad \text{ó} \qquad \overline{\lim}\, (x_n),$$

es el ínfimo del conjunto V de las $v \in \boldsymbol{R}$ tales que hay, cuando más, un número finito de $n \in \boldsymbol{N}$ tal que $v < x_n$.

(b) El **límite inferior** de X, que se designa por medio de

$$\liminf X, \qquad \liminf (x_n), \qquad \text{ó} \qquad \underline{\lim}\, (x_n),$$

es el supremo del conjunto W de las $w \in \boldsymbol{R}$ tales que hay cuando más un número finito de $m \in \boldsymbol{N}$ tal que $x_m < w$.

Aun cuando una sucesión acotada puede no tener un límite, siempre tiene un único límite superior (y un único límite inferior). Esto es claro partiendo del hecho de que el número $v = \sup \{x_n : n \in \boldsymbol{N}\}$ pertenece al conjunto V, mientras que el número $\inf \{x_n : n \in \boldsymbol{N}\} - 1$ es una cota inferior de V.

Existen varias maneras equivalentes, que a menudo son útiles, para definir el límite superior de una sucesión acotada. (Es muy importante que el lector intente comprobar este resultado antes de leer la demostración.)

18.2 TEOREMA. *Si $X = (x_n)$ es una sucesión acotada en $\boldsymbol{R}$, entonces las siguientes afirmaciones son equivalentes para un número real x^*.*

(a) $x^* = \limsup (x_n)$.

(b) Si $\varepsilon > 0$, hay cuando más un número finito de $n \in \boldsymbol{N}$ tal es que $x^* + \varepsilon < x_n$, pero hay un número infinito tal que $x^* - \varepsilon < x_n$.

(c) Si $v_m = \sup \{x_n : n \geq m\}$, entonces $x^* = \inf \{v_m : n \in \boldsymbol{N}\}$.

(d) Si $v_m = \sup \{x_n : n \geq m\}$, entonces $x^* = \lim (v_m)$.

(e) Si L es el conjunto de las $v \in \boldsymbol{R}$ tales que existe una subsucesión de X que converge a v, entonces $x^* = \sup L$.

DEMOSTRACION. Sea $x^* = \limsup (x_n)$ y sea $\varepsilon > 0$. Por la definición 18.1 existe una $v \in V$ con $x^* \leq v \leq x^* + \varepsilon$. Por lo tanto, $x^* + \varepsilon$ también pertenece a V por lo que puede haber cuando más un número finito de $n \in \boldsymbol{N}$

tales que $x^* + \varepsilon < x_n$. Por otro lado, $x^* - \varepsilon$ no está en V, de aquí que haya un número infinito d, $n \in \boldsymbol{N}$ tales que $x^* - \varepsilon < x_n$. Por lo tanto, *(a)* implica *(b)*.

Si *(b)* se cumple, dada $\varepsilon > 0$, entonces para toda m lo suficientemente grande se tiene $v_m \leq x^* + \varepsilon$ por lo tanto, $\inf\{v_m : m \in \boldsymbol{N}\} \leq x^* + \varepsilon$. Pero como hay un número infinito de $n \in \boldsymbol{N}$ tales que $x^* - \varepsilon < x_n$, entonces $x^* - \varepsilon < v_m$ para toda $m \in \boldsymbol{N}$ y por lo tanto $x^* - \varepsilon \leq \inf\{v_m : m \in \boldsymbol{N}\}$. Dado que $\varepsilon > 0$ es arbitrario, se deduce que $x^* = \inf\{v_m : m \in \boldsymbol{N}\}$ y *(c)* se cumple.

Si la sucesión (v_m) se define como en *(c)*, entonces es monótonamente decreciente y por lo tanto $\inf(v_m) = \lim(v_n)$, por lo que *(c)* implica *(d)*.

Ahora, suponga que x^* satisface *(d)* y que $X' = (x_{n_k})$ es una subsucesión convergente de X; dado que $n_k \geq k$ se tiene $x_{n_k} \leq v_k$ y por lo tanto $\lim X' \leq \lim(v_k) = x^*$. Inversamente, obsérvese que existe $n_1 \in \boldsymbol{N}$ tal que $v_1 - 1 < x_{n_1} \leq v_1$. Inductivamente, escoja $n_{k+1} > n_k$ tal que

$$v_k - \frac{1}{k+1} < x_{n_{k+1}} \leq v_k.$$

Dado que $\lim(v_k) = x^*$, se infiere que $x^* = \lim(x_{n_k})$. Por lo tanto, *(d)* implica *(e)*.

Por último, sea $w = \sup L$. Dada $\varepsilon > 0$ puede haber, cuando mucho, un número finito de $n \in \boldsymbol{N}$ con $w + \varepsilon < x_n$ (por el teorema de Bolzano-Weierstrass 16.4). Por lo tanto, $w + \varepsilon \in V$ y $\lim\sup X \leq w + \varepsilon$. Por otro lado, existe una subsucesión X' que converge a un número que excede a $w - \varepsilon$ de donde $w - \varepsilon$ no está en V por lo que $w - \varepsilon \leq \lim\sup X$. Dado que $\varepsilon > 0$ es arbitraria, se deduce que $w = \lim\sup X$. Por lo tanto, *(e)* implica *(a)*.

Q.E.D.

Se puede considerar que ambas caracterizaciones *(d)* y *(e)* justifican al término "límite superior". Existen caracterizaciones correspondientes para el límite inferior de una sucesión acotada, las cuales deberá escribir y demostrar el lector.

En seguida se establecen las propiedades algebraicas básicas del límite superior y del límite inferior de sucesiones acotadas.

18.3 TEOREMA. *Sean* $X = (x_n)$ *y* $Y = (y_n)$ *sucesiones acotadas de números reales. Entonces, las siguientes relaciones son válidas:*

(a) $\lim\inf(x_n) \leq \lim\sup(x_n)$.

(b) Si $c \geq 0$, *entonces* $\lim\inf(cx_n) = c\lim\inf(x_n)$ *y* $\lim\sup(cx_n) = \lim\sup(x_n)$.

(xi) Si $c \leq 0$, *entonces* $\lim\inf(cx_n) = c\lim\sup(x_n)$ *y* $\lim\sup(cx_n) = c\lim\inf(x_n)$.

(c) $\lim\inf(x_n) + \lim\inf(y_n) \leq \lim\inf(x_n + y_n)$.

(d) $\lim\sup(x_n + y_n) \leq \lim\sup(x_n) + \lim\sup(y_n)$.

(e) Si $x_n \leq y_n$ *par toda* n, *entonces* $\lim\inf(x_n) \leq \lim\inf(y_n)$ *y además* $\lim\sup(x_n) \leq \lim\sup(y_n)$.

DEMOSTRACION. *(a)* Si $w < \lim\inf(x_n)$ y $v > \lim\sup(x_n)$, entonces hay un número infinito de $n \in \boldsymbol{N}$ tales que $w \leq x_n$ mientras que sólo hay

un número finito tales que $v < x_n$. Por lo tanto, se debe tener $w \leq v$, que implica *(a)*.

(b) Si $c \geq 0$ entonces la multiplicación por c preserva todas las desigualdades de la forma $w \leq x_n$, etc.

(b') Si $c \leq 0$, entonces la multiplicación por c invierte las desigualdades y convierte al límite superior en límite inferior e inversamente.

La afirmación *(c)* es el dual de *(d)* y se puede deducir directamente de *(d)* o demostrar usando el mismo tipo de argumento. Para demostrar *(d)*, sean $v > \lim\sup(x_n)$ y $u > \lim\sup(y_n)$; por definición, sólo hay un número finito de $n \in \mathbf{N}$ tales que $v < x_n$ y un número finito tales que $u < y_n$. Por lo tanto, sólo puede haber un número finito de n tales que $v + u < x_n + y_n$, probándose que $\lim\sup(x_n + y_n) \leq v + u$. Esto demuestra la afirmación *(d)*.

Se demuestra ahora la segunda afirmación de *(e)*. Si $u > \lim\sup(y_n)$, entonces sólo puede haber un número finito de números naturales n tales que $u < y_n$. Dado que $x_n \leq y_n$, entonces $\lim\sup(x_n) \leq u$ y por lo tanto $\lim\sup(x_n) \leq \lim\sup(y_n)$. Q.E.D.

Cada una de las condiciones equivalentes dadas en el teorema 18.2 se puede usar para demostrar los incisos del teorema 18.3. Se sugiere escribir como ejercicio algunas de estas demostraciones alternativas.

Se podría preguntar si las desigualdades del teorema 18.3 se pueden substituir por igualdades, En general, la respuesta es no. Puesto que, si $X = ((-1)^n)$, entonces $\lim\inf X = -1$ y $\lim\sup X = +1$. Si $Y = ((-1)^{n+1})$, entonces $X + Y = (0)$ de tal manera que

$$\lim\inf X + \lim\inf Y = -2 < 0 = \lim\inf(X + Y),$$

$$\lim\sup(X + Y) = 0 < 2 = \lim\sup X + \lim\sup Y.$$

Se ha visto que el límite inferior y el límite superior existen para cualquier sucesión acotada, no importa si la sucesión es convergente o no. En seguida se demuestra que la existencia $\lim X$ es equivalente a la igualdad de $\lim\inf X$ y $\lim\sup X$.

18.4 LEMA. *Sea X una sucesión acotada de números reales. X es convergente si y sólo si $\lim\inf X = \lim\sup X$, en cuyo caso $\lim X$ es el valor común.*

DEMOSTRACION. Si $x = \lim X$, entonces para cada $\varepsilon > 0$ hay un número natural $N(\varepsilon)$ tal que

$$x - \varepsilon < x_n < x + \varepsilon, \qquad n \geq N(\varepsilon).$$

La segunda desigualdad prueba que $\lim\sup X \leq x + \varepsilon$ y la primera desigualdad prueba que $x - \varepsilon \leq \lim\inf X$. De donde $0 \leq \lim\sup X - \lim\inf X \leq 2\varepsilon$, y puesto que $\varepsilon > 0$ es arbitraria, se tiene la igualdad establecida.

Inversamente, suponga que $x = \lim\inf X = \lim\sup X$. Si $\varepsilon > 0$, a partir del teorema 18.2(b), se deduce que existe un número natural $N_1(\varepsilon)$ tal que si $n \geq N_1(\varepsilon)$, entonces $x_n < x + \varepsilon$. De manera análoga existe un número natural

$N_2(\varepsilon)$ tal que si $n \geq N_2(\varepsilon)$, entonces $x - \varepsilon < x_n$. Sea $N(\varepsilon) = \sup\{N_1(\varepsilon), N_2(\varepsilon)\}$; si $n \geq N(\varepsilon)$, entonces $|x_n - x| < \varepsilon$, probándose que $x = \lim X$.

Sucesiones no acotadas

Algunas veces es conveniente tener definidos el límite superior y el límite inferior para sucesiones arbitrarias (es decir, necesariamente acotadas) en $\boldsymbol{R}$. Para hacer esto es necesario introducir los símbolos $+\infty$ y $-\infty$, se hace la aclaración de que no se consideran como números reales, son simplemente símbolos convenientes.

Si S es un conjunto *no vacío* en $\boldsymbol{R}$ que no está acotado por arriba, se define $\sup S = +\infty$. Si T es un conjunto *no vacío* en $\boldsymbol{R}$ que no está acotado por abajo, se define $\inf T = -\infty$. Todo número real es una cota superior del conjunto vacío $\emptyset$, como se afirmó después de la definición 6.1; entonces se define $\sup \emptyset = -\infty$. Análogamente, todo número real es una cota inferior de $\emptyset$, de manera que se define $\inf \emptyset = +\infty$.

Ahora, sea $X = (x_n)$ una sucesión en $\boldsymbol{R}$ que no está acotada por arriba; entonces el conjunto V de números $v \in \boldsymbol{R}$ tal que hay cuando más un número finito de $n \in \boldsymbol{N}$ tales que $v < x_n$ es vacío. Entonces, el $\inf V = +\infty$. De modo que si $X = (x_n)$ es una sucesión en $\boldsymbol{R}$ que no está acotada por arriba, se tiene

$$\lim\sup (x_n) = +\infty.$$

Análogamente, si $Y = (y_n)$ es una sucesión en $\boldsymbol{R}$ que no está acotada por abajo, se tiene

$$\lim\inf (y_n) = -\infty.$$

Observe que si $X = (x_n)$ es una sucesión en $\boldsymbol{R}$ que no está acotada por arriba, los conjuntos $\{x_n : n \geq m\}$ no están acotados por arriba y se tiene

$$v_m = \sup\{x_n : n \geq m\} = +\infty$$

para toda $m \in \boldsymbol{N}$. Q.E.D.

Límites infinitos

Si $X = (x_n)$ es una sucesión en $\boldsymbol{R}$, se dice que $X = (x_n)$ diverge a $+\infty$, y se escribe $\lim (x_n) = +\infty$, si para toda $\alpha \in \boldsymbol{R}$ hay una $K(\alpha) \in \boldsymbol{N}$ tal que si $n \geq K(\alpha)$; entonces $x_n > \alpha$.

Análogamente, se dice que $X = (x_n)$ diverge a $-\infty$, y se escribe $\lim (x_n) = -\infty$, cuando para toda $\alpha \in \boldsymbol{R}$ hay una $K(\alpha) \in \boldsymbol{N}$ tal que si $n \geq K(\alpha)$ entonces $x_n < \alpha$.

Queda como ejercicio demostrar que $X = (x_n)$ diverge a $+\infty$ si y sólo si

$$\lim\inf (X_n) = \lim\sup (x_n) = +\infty,$$

y que $X=(x_n)$ diverge a $-\infty$ si y sólo si

$$\lim\inf(x_n)=\lim\sup(x_n)=-\infty.$$

Ejercicios

18.A. Determinar el límite superior y el límite inferior de las siguientes sucesiones acotadas en R.

(a) $((-1)^n)$, (b) $((-1)^n/n))$,

(c) $((-1)^n+1/n)$, (d) $(\text{sen } n)$.

18.B. Si $X=(x_n)$ es una sucesión acotada en R, probar que existe una subsucesión de X que converge a lim inf X.

18.C. Formular y demostrar directamente el teorema que corresponde al terorema 18.2 para el límite inferior.

18.D. Dar una demostración directa del teorema 18.3(c)

18.E. Demostrar el teorema 18.3(d) usando 18.2(b) como la definición para el límite superior. Hacer lo mismo usando 18.2(d) y 18.2(e).

18.F. Si $X=(x_n)$ es una sucesión acotada de elementos estrictamente positivos en R demostrar que $\lim\sup(x_n^{1/n})\leq\lim\sup(x_{n+1}/x_n)$.

18.G. Determinar el límite superior y el límite inferior de las siguientes sucesiones en R.

(a) $((-1)^n n)$, (b) $(n\,\text{sen } n)$,

(c) $(n(\text{sen } n)^2)$, (d) $(n\tan n)$.

18.H. Probar que la sucesión $X=(x_n)$ en R diverge a $+\infty$ si y sólo si lim inf $X=+\infty$.

18.I. Probar que lim sup $X=+\infty$ si y sólo si hay una subsucesión X' de X tal que lim $X'=+\infty$.

18.J. Interpretar el teorema 18.3 para sucesiones no acotadas.

Sección 19 Algunas extensiones

En análisis a menudo es importante calcular el "orden de magnitud" de una sucesión o comparar dos sucesiones con respecto a su magnitud. Al hacer esto se descartan los términos que no hacen ninguna "contribución especial". Por ejemplo, si $x_n=2n+17$, entonces, cuando $n\in N$ es grande, la contribución dominante. Se deriva del término $2n$. Si $y_n=n^2-5n$, entonces, cuando $n\in N$ es grande, el término dominante es n^2. A pesar de que los primeros términos de (y_n) son más pequeños que los de (x_n), los términos de esta sucesión, última instancia, crecen más rápidamente que los de (x_n).

Con el objeto de dar mayor precisión a esta idea se introducirán en seguida algunos términos y notaciones (debidos a Landau†) que a menudo son útiles.

†EDMUND (G.H.) LANDAU (1877-1938) fue profesor en Göttingen y es conocido por su trabajo de investigación y sus libros acerca de la teoría de números y análisis. Estos libros se distinguen por lo riguroso y breve del estilo (y su alemán elemental)

19.1 DEFINICION. Sean $X = (x_n)$ y $Y = (y_n)$ sucesiones en $\boldsymbol{R}$ y suponga que $y_n \neq 0$ para toda $n \in \boldsymbol{N}$ lo suficientemente grande. Se dice que X y Y son **equivalentes** y se escribe

$$X \sim Y \qquad \text{ó} \qquad (x_n) \sim (y_n)$$

cuando lim $(x_n/y_n) = 1$. Se dice que X es de un **orden de magnitud menor** que Y y se escribe

$$X = o(Y) \qquad \text{ó} \qquad x_n = o(y_n)$$

cuando lim $(x_n/y_n) = 0$. Se dice que X está dominado por Y y se escribe

$$X = O(Y) \qquad \text{ó} \qquad x_n = O(y_n)$$

cuando la sucesión (x_n/y_n) está acotada.

Es claro que $X \sim Y$ o $X = o(Y)$ implica que $X = O(Y)$. Algunas propiedades de estas notaciones se darán en los ejercicios.

Suma de Cesàro

Ya se ha definido lo que significa la convergencia de una sucesión $X = (x_n)$ en $\boldsymbol{R}^p$ a un elemento x. Sin embargo, es posible adjuntar x a la sucesión X como una especie de "límite generalizado", aun cuando la sucesión X no converja a x en el sentido de la definición 14.3. Existen muchas maneras en que se puede generalizar la idea del límite de una sucesión y darles demasiada importancia, a algunas de ellas, estaría fuera del alcance de este libro. Sin embargo, existe un método que es tanto elemental en su naturaleza como útil en aplicaciones a sucesiones oscilatorias. Dado que es de cierta importancia y la demostración del resultado principal es tipico de muchos argumentos analíticos, se presentará aquí una breve introducción a la teoría de sumabilidad de Cesàro†.

19.2 DEFINICION. Si $X = (x_n)$ es una sucesión de elementos en $\boldsymbol{R}^p$, entonces a la sucesión $S = (\sigma_n)$ definida por

$$\sigma_1 = x_1, \quad \sigma_2 = \frac{x_1 + x_2}{2}, \ldots, \quad \sigma_n = \frac{x_1 + x_2 + \cdots + x_n}{n}, \ldots,$$

se la llama **sucesión de medias aritméticas** de X.

En otras palabras, los elementos de S se encuentran promediando los términos de X. Puesto que este promedio tiende a ablandar fluctuaciones ocasionales en X, es razonable suponer que la sucesión S tiene mayor posibilidad de convergir que la sucesión original X. Cuando la sucesión S de medias aritméticas converge a un elemento y, se dice que la sucesión X es **Cesàro sumable** a y o que y es el **límite** (C,1) de la sucesión X.

Por ejemplo, sea X la sucesión real no convergente $X = (1, 0, 1, 0, \ldots)$; fácilmente se puede ver que si n es un número natural par, entonces $\sigma_n = \frac{1}{2}$ y si n es impar

†ERNESTO CESARO (1859-1906) estudió en Roma y dio clases en Nápoles. Trabajó en geometría y álgebra, así como en análisis.

entonces $\sigma_n = (n+1)/2n$. Dado que $\frac{1}{2} = \lim(\sigma_n)$, la sucesión X es Cesàro sumable a $\frac{1}{2}$, que no es el límite de X pero parece ser el "límite generalizado" más natural que se puede asociar a X.

Es razonable que al generalizar el concepto del límite de una sucesión se requiera que el límite generalizado dé el valor usual del límite siempre que la sucesión sea convergente. En seguida se probará que el método de Cesàro tiene esta propiedad.

19.3 TEOREMA. *Si la sucesión $X = (x_n)$ converge a x, entonces la sucesión $S = (\sigma_n)$ de medias aritméticas también converge a x.*

DEMOSTRACION. Se debe calcular la magnitud de

$$\begin{aligned}\sigma_n - x &= \frac{1}{n}(x_1 + x_2 + \cdots + x_n) - x \\ &= \frac{1}{n}\{(x_1 - x) + (x_2 - x) + \cdots + (x_n - x)\}.\end{aligned} \tag{19.1}$$

Puesto que $x = \lim(x_n)$, entonces dado $\varepsilon > 0$ hay un número natural $N(\varepsilon)$ tal que si $m \geq N(\varepsilon)$, $\|x_m - x\| < \varepsilon$. Además, puesto que la sucesión $X = (x_n)$ es convergente, hay un número real A tal que $\|x_k - x\| < A$ para toda k. Si $n \geq N = N(\varepsilon)$, se descompone la suma del lado derecho de (19.1) en una suma de $k = 1$ to $k = N$ más una suma de $k = N+1$ a $k = n$. Se aplica el cálculo de $\|x_k - x\| < \varepsilon$ a los últimos $n - N$ términos para obtener

$$\|\sigma_n - x\| \leq \frac{NA}{n} + \frac{n-N}{n}\varepsilon \qquad \text{para } n \geq N(\varepsilon).$$

Si n es suficientemente grande, entonces $NA/n < \varepsilon$ y puesto que $(n-N)/n < 1$, se obtiene $\|\sigma_n - x\| < 2\varepsilon$ para n suficientemente grande. Por lo tanto, $x = \lim(\sigma_n)$.

No se insistirá más en la teoría de sumabilidad, pero el lector deberá consultar libros acerca de series divergentes y sumabilidad. Por ejemplo, véase el libro de Knopp que se menciona en la bliografía. Una de las aplicaciones elementales más interesantes de la sumabilidad de Cesàro es el célebre teorema de Fejér que asegura que una función continua se puede reintegrar de su serie de Fourier por medio del proceso de sumabilidad de Cesàro a pesar de que no siempre se puede reintegrar de esta serie por medio de convergencia ordinaria. (Véase el teorema 38.12)

Sucesiones dobles e iteradas

Recalcando, una sucesión en $\boldsymbol{R}^p$ es una función definida en el conjunto de N de números naturales y con rango en $\boldsymbol{R}^p$. Una sucesión doble en $\boldsymbol{R}^p$ es una función X con dominio $\boldsymbol{N} \times \boldsymbol{N}$ que consta de todos los pares ordenados de números naturales y rango en $\boldsymbol{R}^p$. En otras palabras, en cada par ordenado (m, n) de números naturales el valor de la sucesión doble X es un elemento de

$\boldsymbol{R}^p$ que comunmente se designará por medio de x_{mn}. Por lo general se usará un simbolismo como $X = (x_{mn})$ para representar a X, sin embargo, algunas veces es conveniente disponer los elementos en un arreglo tal como

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} & \cdots \\ x_{21} & x_{22} & \cdots & x_{2n} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \end{bmatrix}. \tag{19.2}$$

Observe que en este arreglo el primer índice se refiere a la fila en la que aparece el elemento x_{mn} y el segundo índice se refiere a la columna.

19.4 DEFINICION. Si $X = (x_{mn})$ es una sucesión doble en $\boldsymbol{R}^p$, entonces se dice que un elemento x es un **límite** (o un **límite doble**) de X si para cada número positivo ε hay un número natural $N(\varepsilon)$ tal que para todas $mn \geq N(\varepsilon)$ $\|x_{mn} - x\| < \varepsilon$. En este caso se dice que la sucesión doble **converge** a x y se escribe

$$x = \lim_{mn} (x_{mn}) \qquad \text{ó} \qquad x = \lim X.$$

Gran parte de la teoría elemental de límites de sucesiones, sufriendo muy poco cambio, pasa a sucesiones dobles. En particular, el hecho de que el límite doble esté determinado de manera única (cuando existe) se demuestra exactamente de la misma manera que el teorema 14.5. En forma análoga se pueden definir operaciones algebraicas para sucesiones dobles y obtener resultados exactamente paralelos a los que se analizan en el teorema 15.6. También existe un criterio de Cauchy para la convergencia de una sucesión doble el cual se enunciará pero cuya demostración se deja al lector.

19.5 CRITERIO DE CAUCHY. *Si $X = (x_{mn})$ es una sucesión doble en $\boldsymbol{R}^p$, X es convergente si y sólo si para cada $\varepsilon > 0$ hay un número natural $M(\varepsilon)$ tal que para todas $m, n, r, s \geq M(\varepsilon)$, entonces*

$$\|x_{mn} - x_{rs}\| < \varepsilon.$$

No se insistirá detalladamente en la parte de la teoría de sucesiones dobles que es paralela a la teoría de sucesiones (simples). En cambio, se propone analizar en forma breve la relación que existe entre el límite definido en 19.4 y los límites "iterados".

En primer lugar, se observa que una sucesión doble puede considerarse que da (por lo menos de dos maneras) una *sucesión* de *sucesiones*. Por una parte, se puede considerar a cada renglón del arreglo dado en (19.2) como una sucesión en $\boldsymbol{R}^p$. Así, el primer renglón del arreglo (19.2) da la sucesión $Y_1 = (x_{1n} : n \in \boldsymbol{N}) = (x_{11}, x_{12}, \ldots, x_{1n}, \ldots)$; el segundo renglón en (19.2) da la sucesión $Y_2 = (x_{2n} : n \in \boldsymbol{N})$; etc. Es totalmente razonable tomar en cuenta los

límites de las sucesiones del renglón $Y_1, Y_2, \ldots, Y_m, \ldots$ (siempre que existan dichos límites). Suponiendo que estos límites existen y denotándolos por medio de $y_1, y_2, \ldots, y_m, \ldots$, se obtiene una sucesión de elementos en $\boldsymbol{R}^p$ cuya convergencia bien puede ser examinada. De modo que se está considerando la existencia de $y = \lim (y_m)$. Dado que los elementos y_m están dados por $y_m = \lim Y_m$ en donde $Y_m = (x_{mn} : n \in \boldsymbol{N})$, se designará el límite $y = \lim (y_m)$ (cuando exista) por medio de la expresión

$$y = \lim_m \lim_n (x_{mn}).$$

Se hará referencia a y como un **límite iterado** de la sucesión doble (o con mayor presición como el **límite iterado del renglón** de esta sucesión doble).

Lo que se ha estado haciendo para renglones se puede hacer de igual manera para columnas. De modo que se forman las sucesiones

$$Z_1 = (x_{m1} : m \in \boldsymbol{N}), \qquad Z_2 = (x_{m2} : m \in \boldsymbol{N}),$$

y así sucesivamente. Suponiendo que existan los límites $z_1 = \lim Z_1$, $z_2 = \lim Z_2, \ldots$, se puede considerar entonces $z = \lim (z_n)$. Cuando este úitimo límite existe, se denota por medio de

$$z = \lim_n \lim_m (x_{mn}),$$

y se hace referencia a z como un **límite iterado,** o el **límite iterado de la columna** de la sucesión doble $X = (x_{mn})$.

La primera pregunta que se podría hacer es la siguiente: Si el límite doble de la sucesión $X = (x_{mn})$ existe, ¿entonces existen los límites iterados? La respuesta a esta pregunta puede ser una sorpresa para el lector; es negativa. Para ver esto, sea X una sucesión doble en $\boldsymbol{R}$ dada por $x_{mn} = (-1)^{m+n}(1/m + 1/n)$, fácilmente se puede ver que el límite doble de esta sucesión existe y es 0. Sin embargo, también es fácil probar que ninguna de las sucesiones

$$Y_1 = (x_{1n} : n \in \boldsymbol{N}), \ldots, Y_m = (x_{mn} : n \in \boldsymbol{N}), \ldots$$

tiene límite. Por lo tanto, no es posible que exista ninguno de los límites iterados, ya que ninguno de los límites "internos" existe.

La siguiente pregunta es: Si existen el límite doble y uno de los límites iterados, ¿entonces este límite iterado es igual al límite doble? En este caso la respuesta es afirmativa. De hecho, se establecerá un resultado un poco más fuerte.

19.6 TEOREMA DEL LIMITE DOBLE. *Si el límite doble* $x = \lim_{mn} (x_{mn})$ *existe y si para cada número natural* m *el límite* $y_m = \lim_n (x_{mn})$ *existe, entonces el límite iterado* $\lim_m \lim_n (x_{mn})$ *existe y es igual a* x.

DEMOSTRACION. Por hipótesis, dada $\varepsilon > 0$ hay un número natural $N(\varepsilon)$ tal que si $m, n \geq N(\varepsilon)$, entonces $\|x_{mn} - x\| < \varepsilon$. De nuevo por hipótesis los límites $y_m = \lim_n (x_{mn})$ existen y a partir de la desigualdad anterior y el lema 15.8 se deduce que $\|y_m - x\| \leq \varepsilon$ para toda $m \geq N(\varepsilon)$. Por lo tanto, se concluye que $x = \lim (y_m)$.

El resultado anterior prueba que si el límite doble existe entonces lo único que puede evitar que existan los límites iterados y que sean igual al límite doble es que los límites "internos" no puedan existir. Más claramente, se tiene el siguiente resultado.

19.7 COROLARIO. *Suponga que existe el límite doble y que los límites*

$$y_m = \lim_n (x_{mn}), \qquad z_n = \lim_m (x_{mn})$$

existen para todos los números naturales m, n. Entonces, los límites iterados

$$\lim_m \lim_n (x_{mn}), \qquad \lim_n \lim_m (x_{mn})$$

existen y son iguales al límite doble.

Lo siguiente es preguntar si la existencia e igualdad de los dos límites iterados implica la existencia del límite doble. La respuesta es no. Esto se puede ver analizando la sucesión doble $X = (x_{mn})$ en R definida por $x_{mn} = 1$ cuando $m \neq n$ y por $x_{mn} = 0$ cuando $m = n$. En este caso existen ambos límites iterados y son iguales, pero el límite doble no existe. Sin embargo, con ciertas condiciones adicionales se puede establecer la existencia del límite doble a partir de la existencia de uno de los límites iterados.

19.8 DEFINICION. Para cada número natural m, sea $Y_m = (x_{mn})$ una sucesión en R^p que converge a y_m. Se dice que las sucesiones $\{Y_m : m \in N\}$ son **uniformemente convergentes** si para cada $\varepsilon > 0$ hay un número natural $N(\varepsilon)$ tal que si $n \geq N(\varepsilon)$, entonces $\|x_{mn} - y_m\| < \varepsilon$ para todos los números naturales m.

El lector hará bien en comparar esta definición con la definición 17.4 y observar que son de la misma índole. En cierto modo, para hacer que el teorema 19.10 se prolongue se demuestra que si cada una de las sucesiones Y_m es convergente, entonces la existencia del límite doble implica que las sucesiones $\{Y_m : m \in N\}$ son uniformemente convergentes.

19.9 LEMA. *Si el límite doble de la sucesión doble $X = (x_{mn})$ existe y si para cada número natural m la sucesión $Y_m = (x_{mn} : n \in N)$ es convergnte, entonces esta colección es uniformemente convergente.*

DEMOSTRACION. Dado qu el límite doble existe, dada $\varepsilon > 0$ hay un número natural $N(\varepsilon)$ tal que si $m, n \geq N(\varepsilon)$, entonces $\|x_{mn} - x\| < \varepsilon$. Por

hipótesis, la sucesión $Y_m = (x_{mn} : n \in N)$ converge a un elemento y_m, y aplicando el lema 15.8 se deduce que si $m \geq N(\varepsilon)$,onces $\|y_m - x\| \leq \varepsilon$. Por lo que si $m, n \geq N(\varepsilon)$, se tiene que

$$\|x_{mn} - y_m\| \leq \|x_{mn} - x\| + \|x - y_m\| < 2\varepsilon.$$

Además, para $m = 1, 2, \ldots, N(\varepsilon) - 1$ la sucesión Y_m converge a y_m; por lo tanto, hay un número natural $K(\varepsilon)$ tal que si $n \geq K(\varepsilon)$, entonces

$$\|x_{mn} - y_m\| < \varepsilon, \qquad m = 1, 2, \ldots, N(\varepsilon) - 1.$$

Siendo $M(\varepsilon) = \sup\{N(\varepsilon), K(\varepsilon)\}$, se concluye que si $n \geq M(\varepsilon)$, entonces para cualquier valor de m se tiene

$$\|x_{mn} - y_m\| < 2\varepsilon.$$

Esto ratifica la uniformidad de la convergencia de las sucesiones $\{Y_m : m \in N\}$.

Este último lema muestra que, con la hipótesis de que las sucesiones Y_m convergen, la convergencia uniforme de esta colección de sucesiones es una condición necesaria para la existencia del límite doble. En seguida se da un resultado en el sentido opuesto.

19.10 TEOREMA DEL LIMITE ITERADO. *Suponga que los límites*

$$y_m = \lim_n (x_{mn}), \qquad z_n = \lim_m (x_{mn}),$$

existe para todas $m, n \in N$, *y que la convergencia de una de estas colecciones es uniforme; entonces existen ambos límites iterados y el límite doble, y los tres son iguales.*

DEMOSTRACION. Suponga que la convergencia de la colección $\{Y_m : m \in N\}$ es uniforme. Dada $\varepsilon > 0$, hay un número natural $N(\varepsilon)$ tal que si $n \geq N(\varepsilon)$, entonces

$$\|x_{mn} - y_m\| < \varepsilon \tag{19.3}$$

para todos los números naturales m. Para probar que $\lim (y_m)$ existe, tome un número fijo $q \geq N(\varepsilon)$. Dado que $z_q = \lim (x_{rq} : r \in N)$ existe, se sabe que si $r, s \geq R(\varepsilon, q)$, entonces

$$\|y_r - y_s\| \leq \|y_r - x_{rq}\| + \|x_{rq} - x_{sq}\| + \|x_{sq} - y_s\| < 3\varepsilon.$$

Por lo tanto, (y_r) es una sucesión de Cauchy y converge a un elemento y en R^p. Esto prueba la existencia del límite iterado

$$y = \lim_m (y_m) = \lim_m \lim_n (x_{mn}).$$

En seguida se prueba que el límite doble existe. Dado que $y = \lim(y_m)$, dada $\varepsilon > 0$ hay una $M(\varepsilon)$ tal que si $m \geq M(\varepsilon)$, $\|y_m - y\| < \varepsilon$. Siendo $K(\varepsilon) = \sup\{N(\varepsilon), M(\varepsilon)\}$, de nuevo se usa (19.3) para concluir que si $m, n \geq K(\varepsilon)$, entonces

$$\|x_{mn} - y\| \leq \|x_{mn} - y_m\| + \|y_m - y\| < 2\varepsilon.$$

Esto demuestra que el límite doble existe y es igual a y.

Por último, para demostrar que el otro límite iterado existe y es igual a y, se usa el teorema 19.6 o su corolario.

Se podría deducir que, a pesar de que en la demostración que se acaba de dar se hace uso de la existencia de ambas colecciones de límites y de la uniformidad de una de ellas, la conclusión se puede obtener a partir de la existencia (y uniformidad) de una colección de límites solamente. Se deja al lector investigar si es verdadera o falsa esta deducción.

Ejercicios

19.A. Probar las siguientes relaciones:

(a) $(n^2+2) \sim (n^2-3)$, (b) $(n^2+2) = o(n^3)$,
(c) $((-1)^n n^2) = O(n^2)$, (d) $((-1)^n n^2) = o(n^3)$,
(e) $(\sqrt{n+1} - \sqrt{n}) \sim (1/2\sqrt{n})$, (f) $(\operatorname{sen} n) = O(1)$.

19.B. Sean X, Y y Z sucesiones con elementos distintos de cero. Demostrar que:

(a) $X \sim X$.

(b) Si $X \sim Y$, entonces $Y \sim X$.

(c) Si $X \sim Y$ y $Y \sim Z$, entonces $X \sim Z$.

19.C. Si $X_1 = O(Y)$ y $X_2 = O(Y)$, se deduce que $X_1 \pm X_2 = O(Y)$ y se resume en la "ecuación"

(a) $O(Y) \pm O(Y) = O(Y)$. Dar interpretaciones análogas y demostrar que

(b) $o(Y) \pm o(Y) = o(Y)$.

(c) $c \neq 0$, entonces $o(cY) = o(Y)$ y $O(cY) = O(Y)$.

(d) $O(o(Y)) = o(Y)$, $o(O(Y)) = o(Y)$.

(e) $O(X)O(Y) = O(XY)$, $O(X)o(Y) = o(XY)$, $o(X)o(Y) = o(XY)$.

19.D. Demostrar que $X = o(Y)$ y $Y = o(X)$ no pueden ser válidas simultáneamente. Dar un ejemplo de sucesiones tales que $X = O(Y)$ pero $Y \neq O(X)$.

19.E. Si X es una sucesión monótona en $\boldsymbol{R}$, demostrar que la sucesión de medias aritméticas es monótona.

19.F. Si $X = (x_n)$ es una sucesión en $\boldsymbol{R}$ y (σ_n) es la sucesión de medias aritméticas, entonces $\lim\sup(\sigma_n) \leq \lim\sup(x_n)$. Dar un ejemplo en que se cumpla la igualdad.

19.G. Si $X = (x_n)$ es una sucesión de números reales positivos, ¿es (σ_n) monótonamente creciente?

19.H. Si una sucesión $X = (x_n)$ en $\boldsymbol{R}^p$ es Cesàro sumable, entonces $X = o(n)$. (Sugerencia: $x_n = n\sigma_n - (n-1)\sigma_{n-1}$.)

19.I. Sea X una sucesión monótona en $\boldsymbol{R}$. ¿Es verdad que X es Cesàro sumable si y sólo si es convergente?

19.J. Dar una demostración para el teorema 19.5.

19.K. Considere la existencia de los límites dobles e iterados de las sucesiones dobles (x_{mn}), en donde x_{mn} está dada por

(a) $(-1)^{m+n}$, (b) $\dfrac{1}{m+n}$, (c) $\dfrac{1}{m}+\dfrac{1}{n}$,

(d) $\dfrac{m}{m+n}$, (e) $(-1)^m\left(\dfrac{1}{m}+\dfrac{1}{n}\right)$, (f) $\dfrac{mn}{m^2+n^2}$.

19.L. ¿Es acotada una sucesión doble convergente?

19.M. Si $X=(x_{mn})$ es una sucesión doble convergente de números reales y si para cada $m \in N$, existe el límite $y_m = \lim\sup_n (x_{mn})$ entonces se tiene $\lim_{mn}(x_{mn}) = \lim_m (y_m)$.

19.N. ¿Cuáles de las sucesiones dobles del ejercicio 19.K son tales que la colección $\{Y_m = \lim (x_{mn}) : m \in N\}$ sea uniformemente convergente?

19.O. Sea $X=(x_{mn})$ una sucesión doble acotada en R con la propiedad de que para cada $m \in N$ la sucesión $Y_m = (x_{mn} : n \in N)$ es monótonamente creciente y para cada $n \in N$ la sucesión $Z_n = (x_{mn} : m \in N)$ es monótonamente creciente. ¿Es verdad que existen los límites iterados y que son iguales? ¿existe necesariamente el límite doble?

19.P. Analizar el problema planteado en el último párrafo de esta sección.

IV FUNCIONES CONTINUAS

Se dará principio al estudio de las funciones más importantes en análisis: las funciones continuas. En este capítulo se combinarán los resultados de los capítulos II y III para obtener una colección de teoremas notablemente profundos y útiles.

En la sección 20 se introduce y se analiza el concepto de continuidad. En la sección 21 se introduce la importante clase de funciones lineales. En la sección 22, que es fundamental, se estudian las propiedades de las funciones continuas en conjuntos compactos y conexos; y en la sección 23 se trata el concepto de continuidad uniforme. Los resultados de estas cuatro secciones se emplearán en varias ocasiones en todo lo que resta del libro. Las sucesiones de funciones continuas se estudian en la sección 24, y los límites superior e inferior se estudian en la 25. En lá última sección se ofrecen algunos resultados interesantes e importantes, pero estos resultados no se aplicarán en secciones posteriores.

No se espera que el lector esté familiarizado de antemano con un análisis riguroso de funciones continuas. Sin embargo, en algunos de los ejemplos y de los ejercicios se usan las funciones exponenciales, logarítmicas y trigonométricas con el objeto de proporcionar ciertos ejemplos no triviales.

Sección 20 Propiedades locales de funciones continuas

Se habrá de considerar a f como una función con dominio $D(f)$ contenido en $\mathbf{R}^p$ y con rango $R(f)$ contenido en $\mathbf{R}^q$. En general, no se requerirá que $D(f) = \mathbf{R}^p$ o que $p = q$. Primero se definirá continuidad en términos de vecindades y después se mencionarán algunas condiciones equivalentes.

20.1 DEFINICION. Si $a \in D(f)$, se dice que f es **continua en a** si para toda vecindad V de $f(a)$ existe una vecindad U (dependiendo de V) de a tal que si x es cualquier elemento de $U \cap D(f)$, entonces $f(x)$ es un elemento

i162 (1519x2451x2 pcx)

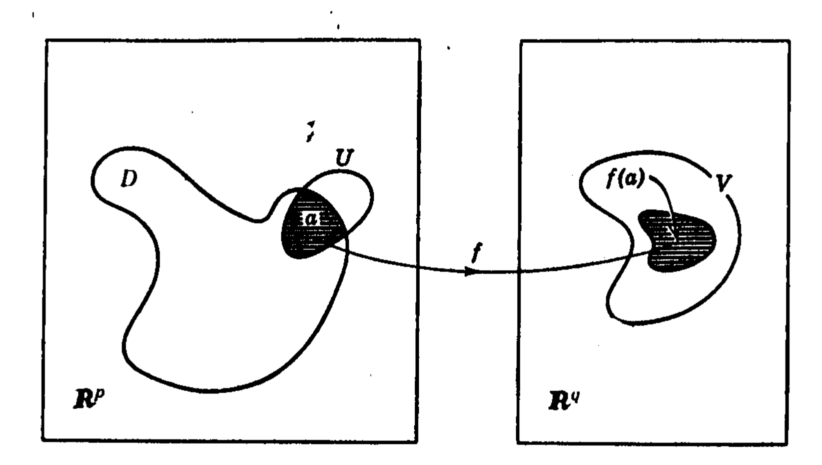


Figura 20.1

de V. (figura 20.1.) Si $A \subseteq D(f)$, se dice que f es **continua en** A siempre que sea continua en todo punto de A.

Algunas veces se dice que una función continua es aquella que "manda puntos vecinales a puntos vecinales" . Esta frase intuitiva se debe eludir si induce a pensar que la imagen de una vecindad de a necesariamente es una vecindad de $f(a)$. (Considerar $x \mapsto |x|$ at $x = 0$.)

En seguida se dan dos afirmaciones equivalentes que se podían haber utilizado como la definición.

20.2 TEOREMA. *Sea a un punto en el dominio $D(f)$ de la función f. Las siguientes afirmaciones son equivalentes:*

(a) f es continua en a.

(b) Si $\varepsilon > 0$, existe un número $\delta(\varepsilon) > 0$ tal que si $x \in D(f)$ y $\|x - a\| < \delta(\varepsilon)$, entonces $\|f(x) - f(a)\| < \varepsilon$.

(c) Si (x_n) es cualquier sucesión de elementos de $D(f)$ que converge a a, entonces la sucesión $(f(x_n))$ converge a $f(a)$.

DEMOSTRACION. Suponga que *(a)* es válido y que $\varepsilon > 0$, entonces la bola $V_\varepsilon = \{y \in R^q : \|y - f(a)\| < \varepsilon\}$ es una vecindad del punto $f(a)$. Por la definición 20.1, hay una vecindad U de a tal que si $x \in U \cap D(f)$, entonces $f(x) \in V_\varepsilon$. Dado que U es una vecindad de a, existe un número real positivo $\delta(\varepsilon)$ tal que la bola abierta con radio $\delta(\varepsilon)$ y centro a está contenida en U. Por lo tanto, la condición *(a)* implica *(b)*.

Suponga que *(b)* es válida y sea (x_n) una sucesión de elementos en $D(f)$ que converge a a. Sea $\varepsilon > 0$ haga referencia a la condición *(b)* para obtener $\delta(\varepsilon) > 0$ con la propiedad establecida en *(b)*. Debido a la convergencia de (x_n) a a, existe un número natural $N(\delta(\varepsilon))$ tal que si $n \geq N(\delta(\varepsilon))$, $\|x_n - a\| < \delta(\varepsilon)$.

Dado que cada $x_n \in D(f)$ de *(b)* se deduce que $\|f(x_n) - f(a)\| < \varepsilon$, demostrándose que *(c)* es válida.

Por último, se razonará indirectamente y demostrará que si la condición *(a)* no es válida entonces la condición *(c)* tampoco lo es. Si *(a)* no se cumple, existe una vecindad V_0 de $f(a)$ tal que para cualquier vecindad U de a existe un elemento x_U que pertenece a $D(f) \cap U$ pero tal que $f(x_U)$ no pertenece a V_0. Para cada número natural n, considere la vecindad U_n de a definida por $U_n = \{x \in \mathbf{R}^p : \|x - a\| < 1/n\}$; por lo establecido en la frase anterior, para cada n en N existe un elemento x_n que pertenece a $D(f) \cap U_n$ pero tal que $f(x_n)$ no pertenece a V_0. La sucesión (x_n) recién construída pertenece a $F(f)$ y converge a a; sin embargo, ninguno de los elementos de la sucesión $(f(x_n))$ pertenece a la vecindad V_0 de $f(a)$. Por lo tanto, se ha construido una sucesión para la cual no es válida la condición *(c)*. Esto demuestra que la parte *(c)* implica *(a)*.

Q.E.D.

El siguiente criterio sobre discontinuidad es una consecuencia de lo que se acaba de hacer.

20.3 CRITERIO SOBRE DISCONTINUIDAD. *La función f no es continua en un punto a en D(f) si y sólo si hay una sucesión* (x_n) *de elementos de D(f) que converge a a pero tal que la sucesión* $(f(x_n))$ *de imágenes no converge a f(a).*

El siguiente resultado es una simple reformulación de la definición. Recuerde, de la definición 2.12 que la **imagen inversa** $f^{-1}(H)$ de un subconjunto H de $\mathbf{R}^q$ bajo f está definida por

$$f^{-1}(H) = \{x \in D(f) : f(x) \in H\}.$$

20.4 TEOREMA. *La función f es continua en un punto a en D(f) si y sólo si para toda vecindad V de f(a) hay una vecindad* V_1 *de a tal que*

$$V_1 \cap D(f) = f^{-1}(V). \qquad (20.1)$$

DEMOSTRACION. Si V_1 es una vecindad de a que satisface esta ecuación, entonces se puede tomar $U = V_1$. Inversamente, si la definición 20.1 se satisface, entonces se puede tomar $V_1 = U \cup f^{-1}(V)$ para obtener la ecuación (20.1).

Q.E.D.

Antes de seguir adelante con la teoría se hará una pausa para dar algunos ejemplos. Por simplicidad, la mayoría de los ejemplos son para el caso en que $\mathbf{R}^p = \mathbf{R}^q = \mathbf{R}$.

20.5 EJEMPLOS. *(a)* Sea $D(f) = \mathbf{R}$ y sea f la función "constante" definida como la función que es igual al número real c para todos los números reales x. Entonces, f es continua en todo punto de $\mathbf{R}$; de hecho, se puede to-

mar la vecindad U, de la definición 20.1, igual a $\boldsymbol{R}$ para cualquier punto a en $D(f)$. Análogamente, la función g definida por

$$\begin{aligned} g(x) &= 1, \qquad 0 \le x \le 1, \\ &= 2, \qquad 2 \le x \le 3, \end{aligned}$$

es continua en cada punto de su dominio.

(b) Sea $D(f)=\boldsymbol{R}$ y sea f la función "identidad" definida como $f(x)=x$, $x \in \boldsymbol{R}$. (figura 20.2) Si a es un número real dado, sea $\varepsilon > 0$ y sea $\delta(\varepsilon)=\varepsilon$. Si $|x-a|<\delta(\varepsilon)$, se tiene $|f(x)-f(a)|=|x-a|<\varepsilon$.

(c) Sea $D(f)=\boldsymbol{R}$ y sea f la función "cuadrática" definida como $f(x)=x^2$, $x \in \boldsymbol{R}$. Sea a elemento de $\boldsymbol{R}$ y tome $\varepsilon>0$; entonces, $|f(x)-f(a)|=|x^2-a^2|=|x-a|\,|x+a|$. Se quiere hacer a esta última expresión menor que ε haciendo a $|x-a|$ lo suficientemente pequeña. Si $a=0$, se elige $\delta(\varepsilon)=\sqrt{\varepsilon}$. Si $a \neq 0$, se desea obtener una cota para $|x+a|$ en una vecindad de a. Por ejemplo, si $|x-a|<|a|$, entonces $0<|x|<2|a|$ y $|x+a| \le |x|+|a| < 3\,|a|$. De donde

(20.2) $$|f(x)-f(a)| \le 3\,|a|\,|x-a|,$$

siempre que $|x-a|<|a|$. Por lo que si se define $\delta(\varepsilon)=\inf\{|a|, \varepsilon/3\,|a|\}$, entonces cuando $|x-a|<\delta(\varepsilon)$, la desigualdad (20.2) es válida y se tiene $|f(x)-f(a)|<\varepsilon$.

(d) Se toma la misma función que en *(c)* pero se usa una técnica un poco distinta. En vez de factorizar x^2-a^2, se escribe como un polinomio en $x-a$. Por lo que

$$x^2-a^2=(x^2-2ax+a^2)+(2ax-2a^2)=(x-a)^2+2a(x-a).$$

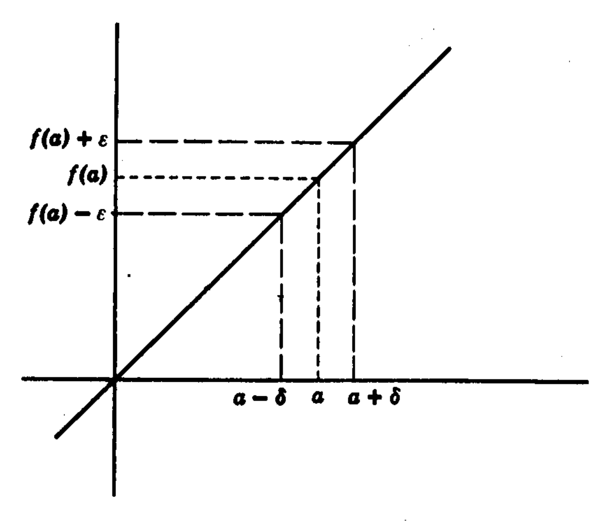


Figura 20.2

Usando la desigualdad del triángulo se obtiene

$$|f(x)-f(a)| \leq |x-a|^2 + 2\,|a|\,|x-a|.$$

Si $\delta \leq 1$ y $|x-a| < \delta$, entonces $|x-a|^2 < \delta^2 \leq \delta$ y el término de la derecha está dominado por $\delta + 2\,|a|\,\delta = \delta(1+2\,|a|)$. Por lo que se escoge

$$\delta(\varepsilon) = \inf\left\{1, \frac{\varepsilon}{1+2\,|a|}\right\}.$$

(e) Considere $D(f) = \{x \in \boldsymbol{R} : x \neq 0\}$ y defina a f como $f(x) = 1/x$, $x \in D(f)$. Si $\in D(f)$, entonces

$$|f(x)-f(a)| = |1/x - 1/a| = \frac{|x-a|}{|ax|}.$$

De nuevo se desea encontrar una cota para el coeficiente $|x-a|$ que sea válida en una vecindad de $a \neq 0$. Observe que si $|x-a| < \frac{1}{2}|a|$, entonces $\frac{1}{2}|a| < |x|$, y se tiene

$$|f(x)-f(a)| \leq \frac{2}{|a|^2}\,|x-a|.$$

Por lo que se debe tomar $\delta(\varepsilon) = \inf\{\frac{1}{2}|a|, \frac{1}{2}\varepsilon\,|a|^2\}$.

(f) Defínase f, para $D(f) = \boldsymbol{R}$ por medio de

$$\begin{aligned} f(x) &= 0, & x &\leq 0, \\ &= 1, & x &> 0. \end{aligned}$$

Se puede ver que f es continua en todos los puntos $a \neq 0$. Se habrá de probar que f no es continua en 0 empleando el criterio de discontinuidad 20.3. De hecho, si $x_n = 1/n$, la sucesión $(f(1/n)) = (1)$ no converge a $f(0)$. (Véase la figura 20.3.)

(g) Sea $D(f) = \boldsymbol{R}$ y sea f la función discontinua de Dirichlet† definida como

$$\begin{aligned} f(x) &= 1, & &\text{si } x \text{ es racional} \\ &= 0, & &\text{si } x \text{ es irracional.} \end{aligned}$$

Si a es un número racional, tome $X = (x_n)$ como una sucesión de números irracionales que convergen a a. (El teorema 6.10 asegura la existencia de dicha sucesión.) Dado que $f(x_n) = 0$ para toda $n \in \boldsymbol{N}$, la sucesión $(f(x_n))$ no converge a $f(a) = 1$ y f no es continua en el número racional a. Por otro lado, si b es un número irracional, existe una sucesión $Y = (y_n)$ de números racionales que converge a b. La sucesión $(f(y_n))$ no con-

† PETER GUSTAV LEJEUNE DIRICHLET (1805-1859) nació en Rhineland y ejerció el magisterio en Berlín durante casi treinta años, antes de irse a Gottingen como sucesor de Gauss Hizo aportaciones fundamentales a la teoría de números y al análisis.

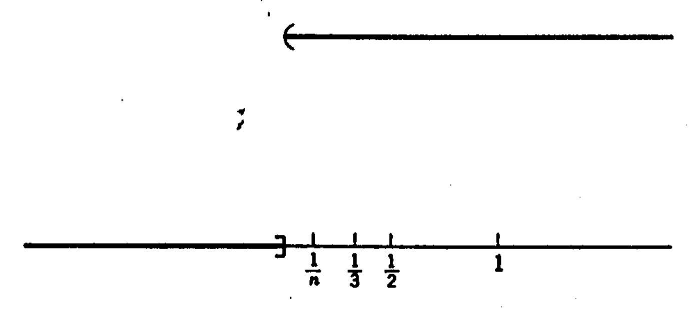


Figura 20.3

verge a $f(b)$ por lo que f no es continua en b. Por lo tanto, la función de Derichlet no es continua en ningún punto.

(h) Sea $D(f) = \{x \in \mathbf{R} : x > 0\}$. Para cualquier número irracional $x > 0$, defínase $f(x) =$ para un número racional de la forma m/n, en donde los números m y n no tienen ningún factor común excepto 1, defínase $f(m/n) = 1/n$. Se habrá de probar que f es continua en todo número irracional en $D(f)$ y discontinua en todo número racional en $D(f)$. La última afirmación resulta de tomar una sucesión de números irracionales que converjan al número racional dado y de usar el criterio de discontinuidad. Sea a un número irracional y $\varepsilon > 0$; entonces existe un número natural n tal que $1/n < \varepsilon$. Si δ se elige tan pequeña que el intervalo $(a - \delta, a + \delta)$ no contenga a ningún número racional con denominador menor que n, entonces se deduce que para x en este intervalo se tiene $|f(x) - f(a)| = |f(x)| \leq 1/n < \varepsilon$. Por lo que f es continua en el número irracional a. Por lo tanto, esta función *es continua precisamente en los puntos irracionales de su dominio.*

(i) En este caso, sea $D(f) = \mathbf{R}^2$ y sea f la función en $\mathbf{R}^2$ con valores en $\mathbf{R}^2$ definida como

$$f(x, y) = (2x + y, x - 3y).$$

Sea (a, b) un punto fijo en $\mathbf{R}^2$; se habrá de probar que f es continua en este punto. Para hacer esto se debe probar que la expresión

$$\|f(x, y) - f(a, b)\| = \{(2x + y - 2a - b)^2 + (x - 3y - a + 3b)^2\}^{1/2}$$

se puede hacer arbitrariamente pequeña escogiendo a (x, y) lo suficientemente cerca de (a, b). Dado que $\{p^2 + q^2\}^{1/2} \leq \sqrt{2} \sup \{|p|, |q|\}$, resulta bastante evidente probar que los términos

$$|2x + y - 2a - b|, \qquad |x - 3y - a + 3b|,$$

se pueden hacer arbitrariamente pequeños eligiendo (x, y) lo suficientemente cerca de (a, b) en $\mathbf{R}^2$. De hecho, por la desigualdad del triángulo,

$$|2x + y - 2a - b| = |2(x - a) + (y - b)| \leq 2\,|x - a| + |y - b|.$$

Ahora, $|x - a| \leq \{(x - a)^2 + (y - b)^2\}^{1/2} = \|(x, y) - (a, b)\|$, y análogamente para $|y - b|$; entonces se tiene

$$|2x+y-2a-b| \leq 3\|(x, y)-(a, b)\|.$$

De modo análogo,

$$|x-3y-a+3b| \leq |x-a|+3|y-b| \leq 4\|(x, y)-(a, b)\|.$$

Por lo tanto, si $\varepsilon > 0$, se puede tomar $\delta(\varepsilon) = \varepsilon/(4\sqrt{2})$ con la seguridad de que si $\|(x, y)-(a, b)\| < \delta(\varepsilon)$, entonces $\|f(x, y)-f(a, b)\| < \varepsilon$, aunque se puede lograr un valor más grande de δ a través de un análisis más refinado (por ejemplo, usando la desigualdad de Schwarz 8.7).

(j) De nuevo sea $D(f) = \mathbf{R}^2$ y defínase a f por medio de

$$f(x, y) = (x^2+y^2, 2xy)$$

Si (a, b) es un punto fijo en $\mathbf{R}^2$,

$$\|f(x, y)-f(a, b)\| = \{(x^2+y^2-a^2-b^2)^2+(2xy-2ab)^2\}^{1/2}.$$

Igual que en (i), se examinan por separado los dos términos de la derecha. Como se podrá ver, se obtendrán cálculos elementales de magnitud. De la desigualdad del triángulo se tiene

$$|x^2+y^2-a^2-b^2| \leq |x^2-a^2|+|y^2-b^2|.$$

Si el punto (x, y) está dentro de una distancia de 1 de (a, b), entonces, $|x| \leq |a|+1$ por lo que $|x+a| \leq 2|a|+1$ y $|y| \leq |b|+1$, de manera que $|y+b| \leq 2|b|+1$. Por lo tanto se tiene

$$\begin{aligned} |x^2+y^2-a^2-b^2| &\leq |x-a|(2|a|+1)+|y-b|(2|b|+1) \\ &\leq 2(|a|+|b|+1)\|(x, y)-(a, b)\|. \end{aligned}$$

De manera análoga se tiene

$$\begin{aligned} |2xy-2ab| = 2|xy-xb+xb-ab| &\leq 2|x||y-b|+2|b||x-a| \\ &\leq 2(|a|+|b|+1)\|(x, y)-(a, b)\|. \end{aligned}$$

Por lo tanto, se establece

$$\delta(\varepsilon) = \inf\left\{1, \frac{\varepsilon}{2\sqrt{2}(|a|+|b|+1)}\right\};$$

si $\|(x, y)-(a, b)\| < \delta(\varepsilon)$, se tiene $\|f(x, y)-f(a, b)\| < \varepsilon$, demostrándose que f es continua en el punto (a, b).

Combinaciones de funciones

El siguiente resultado es consecuencia directa de los teoremas 15.6 y 20.2*(c)*, por lo que no se escribirán los detalles. De manera alternativa, se podría demostrar directamente usando argumentos casi paralelos a los que se emplean en la demostración del teorema 15.6. Se recuerda que si f y g son

funciones con dominios $D(f)$ y $D(g)$ en $\boldsymbol{R}^p$ y rangos en $\boldsymbol{R}^q$, entonces se definen la **suma** $f+g$, la ***diferencia*** $f-g$ y el **producto interno** $f \cdot g$ para cada x en $D(f) \cap D(g)$ por medio de las fórmulas

$$f(x)+g(x), \qquad f(x)-g(x), \qquad f(x) \cdot g(x).$$

Análogamente, si c es un número real y si φ es una función con dominio $D(\varphi)$ en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}$, se definen los productos cf para x en $D(f)$ y φf para x en $D(\varphi) \cap D(f)$ por medio de las fórmulas

$$cf(x), \qquad \varphi(x)f(x).$$

En particular, si $\varphi(x) \neq 0$ para $x \in D_0$, se puede definir el **cociente** f/φ para x en $D(f) \cap D_0$ como

$$f(x)/\varphi(x).$$

Con estas definiciones, se establece el resultado.

20.6 TEOREMA. *Si las funciones f, g, φ son continuas en un punto, entonces las combinaciones algebraicas.*

$$f+g, \qquad f-g, \qquad f \cdot g, \qquad cf, \qquad \varphi f \qquad y \qquad f/\varphi$$

también son continuas en ese punto.

Existe otra combinación algebraica que a menudo resulta útil. Si f está definida para $D(f)$ de $\boldsymbol{R}^p$ a $\boldsymbol{R}$, se define el valor absoluto $|f|$ de f como la función con rango en los números reales $\boldsymbol{R}$ cuyo valor en el punto x en $D(f)$ está dado por $|f(x)|$.

20.7 TEOREMA. *Si f es continua en un punto, entonces $|f|$ también es continua en ese punto.*

DEMOSTRACION. Por la desigualdad del triángulo, se tiene

$$\big|\,|f(x)|-|f(a)|\,\big| \leq |f(x)-f(a)|,$$

de donde el resultado es inmediato. Q.E.D.

Se repetirá el concepto de la composición de dos funciones. Sea f con dominio $D(f)$ en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$ y sea g con dominio $D(g)$ en $\boldsymbol{R}^q$ y rango en $\boldsymbol{R}^r$. En la definición 2.2 se definió la composición $h = g \circ f$ como la función que tiene dominio $D(h) = \{x \in D(f) : f(x) \in D(g)\}$ y en la que para x en $D(h)$ se tiene $h(x) = g[f(x)]$. Por lo que $h = g \circ f$ es una función que transforma a $D(h)$, que es un subconjunto de $D(f) \subseteq \boldsymbol{R}^p$, en un subconjunto de $\boldsymbol{R}^r$. En seguida se demuestra la continuidad de esta función.

20.8 TEOREMA. *Si f es continua en a y g es continua en $b = f(a)$, entonces la composición $g \circ f$ es continua en a.*

DEMOSTRACION. Sea W una vecindad del punto $c = g(b)$. Dado que g es continua en b, existe una vecindad V de b tal que si y pertenece a $V \cap D(g)$, entonces $g(y) \in W$. Como f es continua en a, existe una vecindad U de a tal que si x pertenece a $U \cap D(f)$, entonces $f(x)$ está en V. Por lo tanto, si x pertenece a $U \cap D(g \circ f)$, entonces $f(x)$ está en $V \cap D(g)$ y $g[f(x)]$ pertenece a W. (figura 20.4.) Esto prueba que $h = g \circ f$ es continua en a. Q.E.D.

Ejercicios

20.A. Demostrar que si f está definida para $x \geq 0$ por medio de $f(x) = \sqrt{x}$, entonces f es continua en todo punto de su dominio.

20.B. Demostrar que una "función polinomial", es decir una función f de la forma

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \qquad \text{for } x \in \mathbf{R},$$

es continua en todo punto de $\mathbf{R}$.

20.C. Demostrar que una "función racional" (es decir, el cociente de dos funciones polinomiales) es continua en todo punto en que está definida.

20.D. Usar la desigualdad de Schwarz para demostrar que en el ejemplo 20.5*(i)* se puede tomar $\delta(\varepsilon) = \varepsilon/\sqrt{15}$

20.E. Sea f una función de $\mathbf{R}$ a $\mathbf{R}$ definida por

$$f(x) = x, \quad x \text{ irracional},$$
$$= 1 - x, \quad x \text{ racional}.$$

Demostrar que f es continua en $x = \frac{1}{2}$ y discontinua en los des puntos.

20.F. Sea f continua de $\mathbf{R}$ a $\mathbf{R}$. Demostrar que si $f(x) = 0$ para x racionales, entonces $f(x) = 0$ para toda x en $\mathbf{R}$.

20.G. Sean f y g funciones continuas de $\mathbf{R}$ a $\mathbf{R}$. ¿Es verdad que $f(x) = g(x)$ para $x \in \mathbf{R}$ si y sólo si $f(y) = g(y)$ para todos los números racionales y en $\mathbf{R}$?

20.H. Usar la desigualdad $|\text{sen } x| \leq |x|$ para $x \in \mathbf{R}$ para probar que la función seno es continua en $x = 0$. Usar este hecho, junto con la igualdad

$$\text{sen } x - \text{sen } u = 2 \text{ sen } \tfrac{1}{2}(x - u) \cos \tfrac{1}{2}(x + u),$$

para demostrar que la función seno es continua en cualquier punto de $\mathbf{R}$.

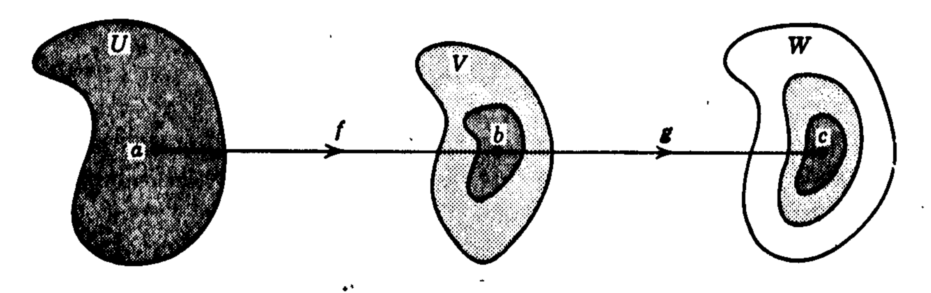


Figura 20.4

20.I. Usando los resultados del ejercicio anterior, demostrar que la función g, definida de $\boldsymbol{R}$ a $\boldsymbol{R}$ como

$$\begin{aligned} g(x) &= x \operatorname{sen}(1/x), \qquad && x \neq 0, \\ &= 0, && x = 0, \end{aligned}$$

es continua en todo punto. Trazar una gráfica de esta función.

20.J. Defínase h, para $x \neq 0$, $x \in \boldsymbol{R}$, como

$$h(x) = \operatorname{sen}(1/x), \qquad x \neq 0.$$

Demostrar que, no importa cómo se defina h en $x = 0$, será discontinua en $x = 0$.

20.K. Defínase a $F : \boldsymbol{R}^2 \to \boldsymbol{R}$ como

$$\begin{aligned} F(x, y) &= x^2 + y^2 \quad && \text{si ambas } x, y \in \boldsymbol{Q}, \\ &= 0 && \text{en los otros casos.} \end{aligned}$$

Determinar los puntos en que F es continua.

20.L. Se dice que una función f de $\boldsymbol{R}$ a $\boldsymbol{R}$ es **aditiva** si satisface

$$f(x + y) = f(x) + f(y)$$

para todas x, $y \in \boldsymbol{R}$. Demostrar que una función aditiva que es continua en $x = 0$ es continua en cualquier punto de $\boldsymbol{R}$. Demostrar que una función aditiva monótona es continua en todo punto.

20.M. Suponga que f es una función aditiva continua en $\boldsymbol{R}$. Si $c = f(1)$, demostrar que $f(x) = cx$ para toda x en $\boldsymbol{R}$. (*Sugerencia:* demostrar primero que si r es un número racional entonces, $f(r) = cr$.)

20.N. Sea $g : \boldsymbol{R} \to \boldsymbol{R}$ tal que satisface la relación

$$g(x + y) = g(x)g(y) \qquad \text{para } x, y \in \boldsymbol{R}.$$

Demostrar si g es continua en $x = 0$, entonces g es continua en todo punto. Además, si $g(a) = 0$ para alguna $a \in \boldsymbol{R}$, $g(x) = 0$ para toda $x \in \boldsymbol{R}$.

20.O. ¿Si $|f|$ es continua en un punto. Es verdad que f también es continua en este punto?

20.P. Sean $f, g : \boldsymbol{R}^p \to \boldsymbol{R}$ continuas en un punto $a \in \boldsymbol{R}^p$ y sean h, k tales que estén definidas de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ por medio de

$$h(x) = \sup\{f(x), g(x)\}, \qquad k(x) = \inf\{f(x), g(x)\}.$$

Demostrar que h y k son continuas en a. (*Sugerencia:* observe que $\sup\{b, c\} = \frac{1}{2}(b + c + |b - c|)$ y $\inf\{b, c\} = \frac{1}{2}(b + c - |b - c|)$.)

20.Q. Si $x \in \boldsymbol{R}$, con frecuencia se define $[x]$ como el máximo entero $n \in \boldsymbol{Z}$ tal que $n \leq x$. A la aplicación $x \mapsto [x]$ se la llama función del **entero máximo.** Trazar las gráficas y determinar los puntos de continuidad de las funciones definidas para $x \in \boldsymbol{R}$ por medio de

(a) $f(x) = [x]$, (b) $g(x) = x - [x]$,
(c) $h(x) = [2 \operatorname{sen} x]$, (d) $k(x) = \operatorname{sen} \frac{1}{2}[x]$.

20.R. Se dice que una función f definida de un intervalo $I \subseteq \boldsymbol{R}$ a $\boldsymbol{R}$ es **creciente** en I si $x \leq x'$, $x, x' \in I$ implica que $f(x) \leq f(x')$. Se dice que es **estrictamente creciente** en I si $x < x'$, $x, x' \in I$ implica que $f(x) < f(x')$. Se pueden dar definiciones análogas para funciones decrecientes y estrictamente decrecientes. Una función que es creciente o bien decreciente en un intervalo se dice que es monótona en ese intervalo.

(a) Si f es creciente en I, f es continua en un punto interior $c \in I$ si y sólo si para toda $\varepsilon > 0$ existen puntos $x_1, x_2 \in I$, $x_1 < c < x_2$, tales que $f(x_2) - f(x_1) < \varepsilon$.

(b) Si f es creciente en I, f es continua en un punto interior $c \in I$ si y sólo si

$$\sup\{f(x) : x < c\} = f(c) = \inf\{f(x) : x > c\}.$$

20.S. Suponga que f es creciente en $I = [a, b]$ en el sentido del ejercicio anterior. Sea

$$j_c = \inf\{f(x) : x > c\} - \sup\{f(x) : x < c\}.$$

Si $j_c > 0$, se dice que f tiene un **salto de** j_c en el punto c.

(a) Si $n \in \boldsymbol{N}$, demostrar que sólo puede haber un conjunto finito de puntos en I en los que f tenga un salto que exceda a $1/n$.

(b) Demostrar que una función creciente puede tener cuando más un conjunto contable de puntos de discontinuidad.

Proyectos

20. α. Sea g una función de $\boldsymbol{R}$ a $\boldsymbol{R}$ que no es idénticamente cero y que satisface la ecuación funcional

$$(*) \qquad g(x+y) = g(x)g(y) \qquad \text{for } x, y \in \boldsymbol{R}.$$

El propósito de este proyecto es demostrar que g debe ser una "función exponencial".

(a) Demostrar que g es continua en todo punto de $\boldsymbol{R}$ si y sólo si es continua en un punto de $\boldsymbol{R}$.

(b) Demostrar que $g(x) > 0$ para toda $x \in \boldsymbol{R}$.

(c) Demostrar que $g(0) = 1$. Si $a = g(1)$, entonces $a > 0$ y $g(r) = a^r$ para toda $r \in \boldsymbol{Q}$.

(d) La función g es estrictamente creciente, es constante, o estrictamente decreciente según sean $g(1) > 1$, $g(1) = 1$, o $0 < g(1) < 1$.

(e) Si $g(x) > 1$ para x en algún intervalo $(0, \delta)$, $\delta > 0$, g es estrictamente creciente y continua en $\boldsymbol{R}$.

(f) Si $a > 0$, existe cuando más una función continua g que satisface $(*)$ tal que $g(1) = a$.

(g) Supóngase que $a > 1$. Haciendo referencia al proyecto 6.β, demostrar que existe una única función continua que satisface $(*)$ tal que $g(1) = a$.

20. β. Sea $P = \{x \in \boldsymbol{R} : x > 0\}$ y sea $h : P \to \boldsymbol{R}$ una función no idénticamente cero que satisfaga la ecuación funcional

$$(\dagger) \qquad h(xy) = h(x) + h(y) \qquad \text{for } x, y \in P.$$

El propósito de este proyecto es demostrar que h debe ser una "función logarítmica".

(a) Demostrar que h es continua en todo punto de P si y sólo si es continua en un punto de P.

(b) Demostrar que h no se puede definir en $x=0$ satisfaciendo (†) para $\{x \in R : x \geq 0\}$.

(c) Demostrar que $h(1)=0$. Si $x>0$ y $r \in Q$, entonces $h(x^r) = rh(x)$.

(d) Demostrar que si $h(x) \geq 0$ en algún intervalo en $\{x \in R : x \geq 1\}$, entonces h es estrictamente creciente y continua en P.

(e) Si h es continua, demostrar que $h(x) \neq 0$ para $x \neq 1$. Además, $h(x)>0$ para $x>1$, o $h(x)<0$ para $x>1$.

(f) Si $b>1$, demostrar que existe, cuando más, una función continua en P que satisface (†) y tal que $h(b)=1$.

(g) Supóngase que $b>1$. Haciendo referencia al proyecto 6. γ, demostrar que existe una función continua única que satisface (†) tal que $h(b)=1$.

Sección 21 Funciones lineales

En la sección anterior se trataron funciones arbitrarias definidas de $\mathbf{R}^p$ a $\mathbf{R}^q$. Antes de seguir con dicho análisis se introducirá una clase de funciones relativamente simple pero muy importante, a saber, las "funciones lineales" que son aplicables en muchísimos casos.

21.1 DEFINICION. Se dice que una función f con dominio $\mathbf{R}^p$ y rango en $\mathbf{R}^q$ es lineal cuando

$$f(ax+by) = af(x)+bf(y) \tag{21.1}$$

para todas a, b en $\mathbf{R}$ y x, y en $\mathbf{R}^p$.

Por inducción se deduce de (21.1) que si $a, b, \dots, c$ son $n \in \mathbf{N}$ números reales y $x, y, \dots, z$ son n elementos de $\mathbf{R}^p$, entonces

$$f(ax+by+\cdots+cz) = af(x)+bf(y)+\cdots+cf(z).$$

Es fácil ver que las funciones de los ejemplos 20.6*(b)* y 20.6*(i)* son funciones lineales para el caso $p=q=1$ y $p=q=2$, respectivamente. De hecho, no es difícil caracterizar en general una función lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$.

21.2 TEOREMA. *Si f es una función lineal con dominio $\mathbf{R}^p$ y rango en $\mathbf{R}^q$, entonces existen pq números reales (c_{ij}), $1 \leq i \leq q$, $1 \leq j \leq p$, tales que si $x=(x_1, x_2, \dots, x_p)$ es cualquier punto en $\mathbf{R}^p$, y si $y=(y_1, y_2, \dots, y_q) = f(x)$ es su imagen bajo f, entonces*

$$\begin{aligned} y_1 &= c_{11}x_1 + c_{12}x_2 + \cdots + c_{1p}x_p, \\ y_2 &= c_{21}x_1 + c_{22}x_2 + \cdots + c_{2p}x_p, \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ y_q &= c_{q1}x_1 + c_{q2}x_2 + \cdots + c_{qp}x_p. \end{aligned} \tag{21.2}$$

i173 (1477x2485x2 pcx)

De manera inversa, si (c_{ij}) *es una colección de pq números reales, entonces la función que asigna a x en* $\mathbf{R}^p$ *el elemento y en* $\mathbf{R}^q$ *según las ecuaciones (21.2), es una función lineal con dominio* $\mathbf{R}^p$ *y rango en* $\mathbf{R}^q$.

DEMOSTRACION. Sean $e_1, e_2, \ldots, e_p$ los elementos de $\mathbf{R}^p$ dados por $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0), \ldots,$ $e_p = (0, 0, \ldots, 1)$. Se analizan las imágenes de estos vectores bajo la función lineal f. Suponga que

$$f(e_1) = (c_{11}, c_{21}, \ldots, c_{q1}),$$
$$f(e_2) = (c_{12}, c_{22}, \ldots, c_{q2}), \tag{21.3}$$
$$\cdots\cdots\cdots\cdots$$
$$f(e_p) = (c_{1p}, c_{2p}, \ldots, c_{qp}).$$

De modo que el número real c_{ij} es la i-ésima coordenada del punto $f(e_j)$.

Un elemento arbitrario $x = (x_1, x_2, \ldots, x_p)$ de $\mathbf{R}^p$ se puede expresar en términos de los vectores $e_1, e_2, \ldots, e_p$; de hecho,

$$x = x_1e_1 + x_2e_2 + \cdots + x_pe_p.$$

Dado que f es lineal, se deduce que

$$f(x) = x_1f(e_1) + x_2f(e_2) + \cdots + x_pf(e_p).$$

Usando las ecuaciones (21.3) se tiene

$$\begin{aligned} f(x) &= x_1(c_{11}, c_{21}, \ldots, c_{q1}) + x_2(c_{12}, c_{22}, \ldots, c_{q2}) \\ &\qquad + \cdots + x_p(c_{1p}, c_{2p}, \ldots, c_{qp}) \\ &= (c_{11}x_1, c_{21}x_1, \ldots, c_{q1}x_1) + (c_{12}x_2, c_{22}x_2, \ldots, c_{q2}x_2) \\ &\qquad + \cdots + (c_{1p}x_p, c_{2p}x_p, \ldots, c_{qp}x_p) \\ &= (c_{11}x_1 + c_{12}x_2 + \cdots + c_{1p}x_p, c_{21}x_1 + c_{22}x_2 + \cdots + c_{2p}x_p, \\ &\qquad \ldots, c_{q1}x_1 + c_{q2}x_2 + \cdots + c_{qp}x_p). \end{aligned}$$

Esto prueba que las coordenadas de $f(x)$ están dadas por las relaciones (21.2), como se aseguraba.

A la inversa, es fácil probar, por cálculos directos, que si las relaciones (21.2) se usan para obtener las coordenadas y_i de y de las coordenadas x_j de x, entonces la función que resulta satisface la relación (21.1) y por lo tanto es lineal. Se omitirá este cálculo debido a que es directo. Q.E.D.

Se debe mencionar que al arreglo rectangular de números

$$\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \cdots & \cdots & \cdots & \cdots \\ c_{q1} & c_{q2} & \cdots & c_{qp} \end{bmatrix}, \tag{21.4}$$

que consta de q renglones y p columnas a menudo se le llama la **matriz** correspondiente a la función lineal f. Hay una correspondencia uno a uno entre las funciones lineales de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ y matrices de $q \times p$ con números reales. Como se ha visto, la acción de f está completamente descrita en términos de su matriz. No será necesario desarrollar la extensa teoría de matrices, pero se hará referencia a la matriz (21.4) como una síntesis de una descripción más elaborada de la función lineal f.

Se demostrará que una función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ es automáticamente continua. Para hacer esto, primero se escribe la desigualdad de Schwarz en la forma

$$|a_1b_1+a_2b_2+\cdots+a_pb_p|^2 \le \{a_1{}^2+a_2{}^2+\cdots+a_p{}^2\}\{b_1{}^2+b_2{}^2+\cdots+b_p{}^2\}.$$

Se aplica esta desigualdad a cada expresión de la ecuación (21.2) para obtener, para $1 \le i \le q$, el cálculo

$$|y_i|^2 \le (|c_{i1}|^2+|c_{i2}|^2+\cdots+|c_{ip}|^2)\,\|x\|^2 = \sum_{j=1}^{p} |c_{ij}|^2\,\|x\|^2.$$

Sumando estas desigualdades se tiene

$$\|y\|^2 \le \left\{\sum_{i=1}^{q}\sum_{j=1}^{p} |c_{ij}|^2\right\}\|x\|^2,$$

de donde se concluye que

$$\|y\| = \|f(x)\| \le \left\{\sum_{i=1}^{q}\sum_{j=1}^{p} |c_{ij}|^2\right\}^{1/2}\|x\|. \tag{21.5}$$

21.3 TEOREMA. *Si f es una función lineal con dominio $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$, entonces existe una constante positiva A tal que si u,v son cualesquiera dos vectores en $\boldsymbol{R}^p$, entonces*

$$\|f(u)-f(v)\| \le A\,\|u-v\|. \tag{21.6}$$

Por lo tanto, una función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ es continua en todo punto.

DEMOSTRACION. Al obtener la fórmula (21.5) se vio que existe una constante A tal que si x es cualquier elemento de $\boldsymbol{R}^p$ entonces $\|f(x)\| \le A\,\|x\|$. Ahora, sea $x=u-v$ usando la linealidad de f se obtiene $f(x)=f(u-v)=f(u)-f(v)$. Por lo tanto, se obtiene la fórmula (21.6). Es claro que esta relación implica la continuidad de f ya que se puede hacer $\|f(u)-f(v)\| < \varepsilon$ tomando $\|u-v\| < \varepsilon/A$ if $A>0$. Q.E.D.

Queda como ejercicio demostrar que si f y g son funciones lineales de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$, entonces $f+g$ es una función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$. Análogamente, si $c \in \boldsymbol{R}$, entonces cf es una función lineal. Se deja al lector demostrar que la colección $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ de todas las funciones lineales de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ es un espacio vectorial bajo estas operaciones vectoriales. En los ejercicios se mostrará cómo definir una norma en este espacio vectorial.

Ejercicios

21.A. Demostrar que $f: \mathbf{R}^p \to \mathbf{R}^q$ es una función lineal si y sólo si $f(ax) = af(x)$ y $f(x+y) = f(x) + f(y)$ para toda $a \in \mathbf{R}$ y todas $x, y \in \mathbf{R}^p$.

21.B. Si f es una función lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$, demostrar que las columnas de la representación matricial (21.4) de f indican los elementos en $\mathbf{R}^q$ a los que los elementos $e_1 = (1, 0, \ldots, 0)$, $e_2 = (0, 1, \ldots, 0), \ldots, e_p = (0, 0, \ldots, 1)$ de $\mathbf{R}^p$ se aplican por medio de f.

21.C. Sea f una función lineal de $\mathbf{R}^2$ a $\mathbf{R}^3$ que manda a los elementos $e_1 = (1, 0)$, $e_2 = (0, 1)$ de $\mathbf{R}^2$ hacia los vectores $f(e_1) = (2, 1, 0)$, $f(e_2) = (1, 0, -0)$ de $\mathbf{R}^3$. Dar la representación matricial de f. ¿Cuáles vectores en $\mathbf{R}^3$ son las imágenes bajo f de los elementos (2,0), (1,1) y (1,3)?

21.D. Si f denota a la función lineal del ejercicio 21.C, demostrar que no todo vector en $\mathbf{R}^3$ es la imagen bajo f de algún vector en $\mathbf{R}^2$.

21.E. Sea g cualquier función lineal de $\mathbf{R}^2$ a $\mathbf{R}^3$. Demostrar que no todo elemento de $\mathbf{R}^3$ es la imagen bajo g de algún vector en $\mathbf{R}^2$.

21.F. Sea h cualquier función lineal de $\mathbf{R}^3$ a $\mathbf{R}^2$. Demostrar que existen vectores distintos de cero en $\mathbf{R}^3$ que se aplican hacia el vector cero de $\mathbf{R}^2$ por h.

21.G. Sea f una función lineal de $\mathbf{R}^2$ a $\mathbf{R}^2$ y sea la matriz que representa a f la siguiente

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

Demostrar que $f(x) \neq 0$ cuando $x \neq 0$ si y sólo si $\Delta = ad - bc \neq 0$.

21.H. Sea f como en el ejercicio 21.G. Demostrar que f aplica $\mathbf{R}^2$ sobre $\mathbf{R}^2$ si y sólo si $\Delta = ad - bc \neq 0$. Demostrar que si $\Delta \neq 0$, entonces la función inversa f^{-1} es lineal y la matriz que la representa es la siguiente

$$\begin{bmatrix} d/\Delta & -b/\Delta \\ -c/\Delta & a/\Delta \end{bmatrix}$$

21.I. Sea g una función lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$. Demostrar que g es uno a uno si y sólo si $g(x) = 0$ implica que $x = 0$.

21.J. Si h es una función lineal uno a uno de $\mathbf{R}^p$ sobre $\mathbf{R}^p$, demostrar que la inversa h^{-1} es una función lineal de $\mathbf{R}^p$ sobre $\mathbf{R}^p$.

21.K. Demostrar que la suma y la composición de dos funciones lineales son funciones lineales.

21.L. Si f es una aplicación lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$, definir

$$\|f\|_{pq} = \sup\{\|f(x)\| : x \in \mathbf{R}^p, \|x\| \leq 1\}.$$

Demostrar que la aplicación $f \mapsto \|f\|_{pq}$ define una norma en el espacio vectorial $\mathscr{L}(\mathbf{R}^p, \mathbf{R}^q)$ de todas las funciones lineales de $\mathbf{R}^p$ a $\mathbf{R}^q$. Demostrar que $\|f(x)\| \leq \|f\|_{pq} \|x\|$ para toda $x \in \mathbf{R}^p$.

21.M. Si f es una aplicación lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$, definir

$$M(f) = \inf\{M > 0 : \|f(x)\| \leq M \|x\|, x \in \mathbf{R}^p\}.$$

Demostrar que $M(f) = \|f\|_{pq}$.

21.N. Si f y g están en $\mathcal{L}(\mathbf{R}^p, \mathbf{R}^p)$ demostrar que $f \circ g$ también está en $\mathcal{L}(\mathbf{R}^p, \mathbf{R}^p)$ y que $\|f \circ g\|_{pp} \leq \|f\|_{pp} \|g\|_{pp}$. Demostrar que la desigualdad puede ser estricta para ciertas f y g.

21.O. Dar un ejemplo de una aplicación lineal f en $\mathcal{L}(\mathbf{R}^p, \mathbf{R}^q)$ con representación matricial $[c_{ij}]$ en donde se tenga

$$\|f\|_{pq} < \left\{ \sum_{i=1}^{q} \sum_{j=1}^{p} c_{ij}^2 \right\}^{1/2}$$

21.P. Si (21.4) da la matriz para f, demostrar que $|c_{ij}| \leq \|f\|_{pq}$ para todas i, j.

Sección 22 Propiedades globales de funciones continuas

En la sección 20 se consideró la continuidad "local"; es decir, se estaba tratando la continuidad en un punto. En esta sección se establecerán algunas propiedades más profundas acerca de funciones continuas. Aquí se estará tratando la continuidad "global" en el sentido de que se supondrá que las funciones son continuas en todo punto de sus dominios.

A menos que se especifique lo contrario, f designará a una función con dominio $D(f)$ contenido en $\mathbf{R}^p$ y con rango en $\mathbf{R}^q$. Se recuerda que si B es subconjunto del espacio de rango $\mathbf{R}^q$, la **imágen inversa** de B bajo f es el conjunto

$$f^{-1}(B) = \{x \in D(f) : f(x) \in B\}.$$

Observe que $f^{-1}(B)$ de manera automática es un subconjunto de $D(f)$ aun cuando B no necesariamente sea un subconjunto del rango de f.

En cursos de topología, en que se trata más continuidad global que continuidad local, el siguiente resultado a menudo se toma como la definición de continuidad (global). Su importancia muy pronto será evidente.

22.1 TEOREMA DE CONTINUIDAD GLOBAL. *Las siguientes afirmaciones son equivalentes:*

(a) *f es continua en su dominio $D(f)$.*

(b) *Si G es cualquier conjunto abierto en $\mathbf{R}^q$, entonces existe un conjunto abierto G_1 en $\mathbf{R}^p$ tal que $G_1 \cap D(f) = f^{-1}(G)$.*

(c) *Si H es cualquier conjunto cerrado en $\mathbf{R}^q$, entonces existe un conjunto cerrado H_1 en $\mathbf{R}^p$ tal que $H_1 \cap D(f) = f^{-1}(H)$.*

DEMOSTRACION. Primero se supondrá que (a) es válido y que G es un subconjunto abierto de $\mathbf{R}^q$. Si a pertenece a $f^{-1}(G)$, entonces, como G es una vecindad de $f(a)$, de la continuidad de f en a se infiere que hay un conjunto abierto $U(a)$ tal que si $x \in D(f) \cap U(a)$, entonces $f(x) \in G$. Elíjase $U(a)$ para cada a en $f^{-1}(G)$ y sea G_1 la unión de los conjuntos $U(a)$. Por el teo-

rema 9.3(c), el conjunto G_1 es abierto y es claro que $G_1 \cap D(f) = f^{-1}(G)$. Por lo tanto, (a) implica (b).

Se demostrará ahora que (b) implica (a). Si a es un punto arbitrario de $D(f)$ y G es una vecindad abierta de $f(a)$, la condición (b) implica que existe un conjunto abierto G_1 en $\boldsymbol{R}^p$ tal que $G_1 \cap D(f) = f^{-1}(G)$. Dado que $f(a) \in G$, se deduce que $a \in G_1$, por lo que G_1 es una vecindad de a. Si $x \in G_1 \cap D(f)$, $f(x) \in G$ por lo que f es continua en a. Esto prueba que la condición (b) implica (a).

Ahora se demuestra la equivalencia de las condiciones (b) y (c). Primero, se observa que si B es cualquier subconjunto de $\boldsymbol{R}^q$ y si $C = \boldsymbol{R}^q \setminus B$, se tiene $f^{-1}(B) \cap f^{-1}(C) = \emptyset$ y

$$(22.1) \qquad D(f) = f^{-1}(B) \cup f^{-1}(C).$$

Si B_1 es un subconjunto de $\boldsymbol{R}^p$ tal que $B_1 \cap D(f) = f^{-1}(B)$ y $C_1 = \boldsymbol{R}^p \setminus B_1$, entonces $C_1 \cap f^{-1}(B) = \emptyset$ y

$$(22.2) \qquad D(f) = (B_1 \cap D(f)) \cup (C_1 \cap D(f)) = f^{-1}(B) \cup (C_1 \cap D(f)).$$

Las fórmulas (22.1) y (22.2) son dos representaciones de $D(f)$ como la unión de $f^{-1}(B)$ con otro conjunto con el que no tiene puntos en común. Por lo tanto, se tiene $C_1 \cap D(f) = f^{-1}(C)$.

Suponga que (b) es válido y que H es cerrado en $\boldsymbol{R}^q$. Aplique el argumento dado anteriormente en el caso en que $B = \boldsymbol{R}^q \setminus H$ y $C = H$. Entonces, B y B_1 son conjuntos abiertos en $\boldsymbol{R}^q$ y $\boldsymbol{R}^p$, respectivamente, de manera que $C_1 = \boldsymbol{R}^q \setminus B_1$ es cerrado en $\boldsymbol{R}^p$. Esto prueba que (b) implica (c).

Para ver que (c) implica (b), se usa el argumento anterior con $B = \boldsymbol{R}^p \setminus G$, en donde G es un conjunto abierto en $\boldsymbol{R}^q$. Q.E.D.

En el caso en que $D(f) = \boldsymbol{R}^p$, el resultado anterior se simplifica hasta cierto punto.

22.2 COROLARIO. *Sea f una función definida en todo $\boldsymbol{R}^p$ y con rango en $\boldsymbol{R}^q$. Las siguientes afirmaciones son equivalentes:*

(a) *f es continua en $\boldsymbol{R}^p$;*

(b) *si G es abierto en $\boldsymbol{R}^q$, entonces $f^{-1}(G)$ es abierto en $\boldsymbol{R}^p$;*

(c) *si H es cerrado en $\boldsymbol{R}^q$, entonces $f^{-1}(H)$ es cerrado en $\boldsymbol{R}^p$.*

Se debe hacer énfasis en que el teorema de continuidad global 22.1 *no* dice que si f es continua y G un conjunto abierto en $\boldsymbol{R}^p$, entonces la imagen directa $f(G) = \{f(x) : x \in G\}$ sea abierta en $\boldsymbol{R}^q$. En general, una función continua no necesariamente manda conjuntos abiertos a conjuntos abiertos o conjuntos cerrados a conjuntos cerrados. Por ejemplo, la función f de $\boldsymbol{R}$ a $\boldsymbol{R}$ definida como

$$f(x) = \frac{1}{1+x^2}$$

es continua en $\boldsymbol{R}$.[En efecto, en los ejemplos 20.5(a) y (c) se vió que las funciones $f_1(x) = 1$, y $f_2(x) = x^2$, para $x \in \boldsymbol{R}$, son continuas en todo punto. Del teorema 15.6 se deduce que

$$f_3(x) = 1 + x^2, \qquad x \in \boldsymbol{R},$$

es continua en todo punto, y dado que f_3 nunca se hace cero, este mismo teorema implica que la función f recién dada es continua en $\boldsymbol{R}$.] Si G es el conjunto abierto $G = (-1, 1)$, entonces $f(G) = (\frac{1}{2}, 1]$, que no es abierto en $\boldsymbol{R}$. Análogamente, si H es el conjunto cerrado $H = \{x \in \boldsymbol{R} : x \geq 1\}$, entonces $f(H) = (0, \frac{1}{2}]$, que no es cerrado en $\boldsymbol{R}$. De manera similar, la función f aplica al conjunto $\boldsymbol{R}$, que es abierto y cerrado en $\boldsymbol{R}$, hacia el conjunto $f(\boldsymbol{R}) = (0, 1]$ que no es ni abierto ni cerrado en $\boldsymbol{R}$.

Lo que nos dicen estas afirmaciones es que la propiedad de un conjunto de ser cerrado o abierto no necesariamente se preserva con la aplicación de una función continua. Sin embargo, existen propiedades importantes de conjunto que se preservan con aplicaciones continuas. Por ejemplo, se va a demostrar que las propiedades de conexidad y compacidad de conjuntos tienen esta característica.

Conservación de conexidad

Por la definición 12.1, recuerde que un conjunto H en $\boldsymbol{R}^p$ es descone si existen conjuntos abiertos A, B en $\boldsymbol{R}^p$ tales que $A \cap H$ y $B \cap H$ son conjuntos ajenos no vacíos cuya unión es H. Un conjunto es conexo si no es disconexo.

22.3 CONSERVACION DE CONEXIDAD. *Si $H \subseteq D(f)$ es conexo en $\boldsymbol{R}^p$ y f es continua en H, entonces $f(H)$ es conexa en $\boldsymbol{R}^q$.*

DEMOSTRACION. Sea h la restricción de f al conjunto H de tal manera que $D(h) = H$ y $h(x) = f(x)$ para toda $x \in H$. Observe que $f(H) = h(H)$ y que h es continua en H.

Si $f(H) = h(H)$ es disconexo en $\boldsymbol{R}^q$, entonces existen conjuntos abiertos A, B en $\boldsymbol{R}^q$ tales que $A \cap h(H)$ y $B \cap h(H)$ son conjuntos ajenos no vacíos cuya unión es $h(H)$. Por el teorema de continuidad global 22.1, existen conjuntos abiertos A_1, B_1 en $\boldsymbol{R}^p$ tales que

$$A_1 \cap H = h^{-1}(A), \qquad B_1 \cap H = h^{-1}(B).$$

Estas intersecciones son no vacías y se deduce que son ajenas de que los conjuntos $A \cap h(H)$ y $B \cap h(H)$. son ajenos. El supuesto de que la unión de $A \cap h(H)$ y $B \cap h(H)$ sea $h(H)$ implica que la unión de $A_1 \cap H$ y $B_1 \cap H$ es H. Por lo tanto, la disconexión de $f(H) = h(H)$ implica la disconexión de H.

Q.E.D.

La misma palabra "continuo" indica que no hay "interrupciones" bruscas en la gráfica de la función, por lo que el siguiente resultado no es de ninguna manera inesperado. Sin embargo, se invita al lector a que intente proporcionar una demostración distinta para este teorema y así se dé cuenta de su profundidad.

22.4 TEOREMA DEL VALOR INTERMEDIO DE BOLZANO. *Sea $H \subseteq D(f)$ un subconjunto conexo de $\mathbf{R}^p$ y sea f continua en H y con valores en* $\mathbf{R}$. *Si k es cualquier número real que satisface*

$$\inf\{f(x): x \in H\} < k < \sup\{f(x): x \in H\},$$

entonces hay cuando menos un punto de H en donde f toma el valor k.

DEMOSTRACION. Si $k \notin f(H)$, entonces los conjuntos $A = \{t \in \mathbf{R} : t < k\}$, $B = \{t \in \mathbf{R} : t > k\}$ forman una disconexión de $f(H)$, contradiciéndose el teorema anterior. Q.E.D.

Conservación de compacidad

En seguida se demuestra que la importante propiedad de compacidad se conserva con la aplicación continua. Recuerde que es consecuencia del teorema de Heine-Borel 11.3 que un subconjunto K de $\mathbf{R}^p$ sea compacto si y sólo si es cerrado y acotado en $\mathbf{R}^p$. De modo que el siguiente resultado se puede expresar diciendo que si K es cerrado y acotado en $\mathbf{R}^p$ y si f es continua en K y con rango en $\mathbf{R}^q$, entonces $f(K)$ es cerrado y acotado en $\mathbf{R}^q$.

22.5 CONSERVACION DE COMPACIDAD. *Si $K \subseteq D(f)$ es compacto y f es continua en K, entonces $f(K)$ es compacto.*

PRIMERA DEMOSTRACION. Se supone que K es cerrado y acotado en $\mathbf{R}^p$; se habrá de probar que $f(K)$ es cerrado y acotado en $\mathbf{R}^q$. Si $f(K)$ no es acotado, para cada $n \in \mathbf{N}$ existe un punto x_n en K con $\|f(x_n)\| \geq n$. Dado que K es acotado, la sucesión $X = (x_n)$ es acotada, de donde, por el teorema de Bolzano-Weierstrass 16.4, se deduce que hay una subsucesión de X que converge a un elemento x. Dado que $x_n \in K$ para $n \in \mathbf{N}$, el punto x pertenece al conjunto cerrado K. Por lo tanto, f es continua en x y f está acotada por $\|f(x)\| + 1$ en una vecindad de x. Dado que esto contradice al supuesto de que $\|f(x_n)\| \geq n$, el conjunto $f(K)$ es acotado.

Se habrá de probar que $f(K)$ es cerrado demostrando que cualquier punto de acumulación y de $f(K)$ debe estar contenido en este conjunto. De hecho, si n es un número natural, existe un punto z_n en K tal que $\|f(z_n) - y\| < 1/n$. Por el teorema de Bolzano-Weierstrass 16.4, la sucesión $Z = (z_n)$ tiene una subsucesión $Z' = (z_{n(k)})$ que converge a un elemento z. Dado que K es cerrado, $z \in K$ y f es continua en z. Por lo tanto,

$$f(z) = \lim_k (f(z_{n(k)})) = y,$$

lo cual demuestra que y pertenece a $f(K)$. Por lo tanto, $f(K)$ es cerrado.

SEGUNDA DEMOSTRACION. Restringiendo f a K se puede suponer que $D(f) = K$. Suponga ahora que $\mathscr{G} = \{G_\alpha\}$ es una familia de conjuntos abiertos en $\boldsymbol{R}^q$ cuya unión contiene a $f(K)$. Por el teorema de continuidad global 22.1, para cada conjunto G_α en $\mathscr{G}$ hay un subconjunto abierto C_α de $\boldsymbol{R}^p$ tal que $C_\alpha \cap D = f^{-1}(G_\alpha)$. La familia $\mathscr{C} = \{C_\alpha\}$ consta de subconjuntos abiertos de $\boldsymbol{R}^p$; se asegura que la unión de estos conjuntos contiene a K, ya que si $x \in K$, entonces $f(x)$ está contenido en $f(K)$; por lo tanto, $f(x)$ pertenece a algun conjunto G_α y, por construcción, x pertenece al conjunto correspondiente C_α. Dado que K es compacto, está contenido en la unión de un número finito de conjuntos en $\mathscr{C}$ y su imagen $f(K)$ está contenida en la unión del número finito de conjuntos correspondientes en $\mathscr{G}$. Dado que esto es válido para una familia arbitraria $\mathscr{G}$ de conjuntos abiertos que cubren a $f(K)$, el conjunto $f(K)$ es compacto en $\boldsymbol{R}^q$. Q.E.D.

Cuando el rango de la función es $\boldsymbol{R}$, el siguiente teorema se puede expresar de otra manera diciendo que *una función continua de valor real en un conjunto compacto alcanza sus valores máximo y mínimo.*

22.6 TEOREMA DEL VALOR MAXIMO Y MINIMO. *Sea $K \subseteq D(f)$ compacto en $\boldsymbol{R}^p$ y sea f función continua de valor real. Entonces existen puntos x^* en K tales que*

$$f(x^*) = \sup\{f(x) : x \in K\}, \qquad f(x_*) = \inf\{f(x) : x \in K\}.$$

PRIMERA DEMOSTRACION. Dado que K es compacto en $\boldsymbol{R}^p$, del teorema anterior se deduce que $f(K)$ está acotado en $\boldsymbol{R}$. Sean $M = \sup f(K)$ y (x_n) una sucesión en K tal que

$$f(x_n) \geq M - 1/n, \qquad n \in \boldsymbol{N}.$$

Por el teorema de Bolzano-Weierstrass 16.4, alguna subsucesión $(x_{n(k)})$ converge a un límite $x^* \in K$. Dado que f es continua en x^*, se debe tener $f(x^*) = \lim(f(x_{n(k)})) = M$. La demostración de la existencia de x_* es análoga.

SEGUNDA DEMOSTRACION. Restringiendo f a K, se puede suponer que $D(f) = K$. Se fija $M = \sup f(K)$. Entonces, para cada $n \in \boldsymbol{N}$, sea $G_n = \{u \in \boldsymbol{R} : u < M - 1/n\}$. Dado que G_n es abierto, del teorema de continuidad global 22.1 se deduce que existe un conjunto abierto C_n en $\boldsymbol{R}^p$ tal que

$$C_n \cap K = \{x \in K : f(x) < M - 1/n\}.$$

Si no se alcanza el valor M entonces la unión de la familia $\mathscr{C} = \{C_n\}$ de conjuntos abiertos contiene a todo K. Dado que K es compacto y la familia $\{C_n \cap K\}$ es creciente, existe una $r \in \boldsymbol{N}$ tal que $K \subseteq C_r$ pero entonces se tiene $f(x) < M - 1/r$ para toda $x \in K$, contrario al hecho de que $M = \sup f(K)$.

Q.E.D.

Si f tiene rango en $\boldsymbol{R}^q$ con $q > 1$, el siguiente corolario algunas veces es útil.

22.7 COROLARIO. *Sea f una función de $D(f) \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^q$ y sea $K \subseteq D(f)$ compacto. Si f es continua en K, existen puntos x^* en K tales que*

$$\|f(x^*)\| = \sup\{\|f(x)\| : x \in K\}, \qquad \|f(x_*)\| = \inf\{\|f(x)\| : x \in K\}.$$

Del teorema 21.2 se infiere que si $f: \boldsymbol{R}^p \to \boldsymbol{R}^q$ es lineal, entonces existe una constante $M > 0$ tal que $\|f(x)\| \leq M\,\|x\|$ para toda $x \in \boldsymbol{R}^p$. Sin embargo, no siempre ocurre que exista una constante $m > 0$ tal que $\|f(x)\| \geq m\,\|x\|$ para toda $x \in \boldsymbol{R}^p$. En seguida se demuestra que este es el caso si y sólo si f es una función lineal inyectiva.

22.8 COROLARIO. *Sea $f: \boldsymbol{R}^p \to \boldsymbol{R}^q$ una función lineal. Entonces f es inyectiva si y sólo si existe $m > 0$ tal que $\|f(x)\| \geq m\,\|x\|$ para toda $x \in \boldsymbol{R}^p$.*

DEMOSTRACION. Suponga que f es inyectiva y sea $S = \{x \in \boldsymbol{R}^p : \|x\| = 1\}$ es la esfera compacta unitaria en $\boldsymbol{R}^p$. Por el corolario 22.7, existe $x_* \in S$ tal que $\|f(x_*)\| = m = \inf\{\|f(x)\| : x \in S\}$. Dado que f es inyectiva, $m = \|f(x_*)\| > 0$. Por lo tanto, $\|f(x)\| \geq m > 0$ para toda $x \in S$. Ahora, si $u \in \boldsymbol{R}^p$, $u \neq 0$, entonces $u/\|u\|$ pertenece a S y por la línealidad de f se tiene

$$\frac{1}{\|u\|}\|f(u)\| = \left\|f\left(\frac{u}{\|u\|}\right)\right\| \geq m,$$

por lo que se deduce que $\|f(u)\| \geq m\|u\|$ para toda $u \in \boldsymbol{R}^p$ (ya que el resultado es trivial para $u = 0$).

De manera inversa suponga que $\|f(x)\| \leq m\,\|x\|$ para toda $x \in \boldsymbol{R}^p$. Si $f(x_1) = f(x_2)$, se tiene

$$0 = \|f(x_1) - f(x_2)\| = \|f(x_1 - x_2)\| \geq m\,\|x_1 - x_2\|,$$

que implica $x_1 = x_2$. Por lo tanto, f es inyectiva. Q.E.D.

Una de las consecuencias más notables del teorema 22.5 es que si f es continua e inyectiva en un dominio compacto entonces la función inversa f^{-1} automáticamente es continua.

22.9 CONTINUIDAD DE LA FUNCION INVERSA. *Sea K un subconjunto compacto de $\boldsymbol{R}^p$ y sea f una función continua inyectiva con do-*

minio K y rango $f(K)$ en $\mathbf{R}^q$. Entonces la función inversa es continua con dominio $f(K)$ y rango K.

DEMOSTRACION. Observe que como K es compacto, el teorema 22.5 implica que $f(K)$ es compacto y por lo tanto cerrado. Dado que por hipótesis f es inyectiva, la función inversa $g = f^{-1}$ está definida. Sea H cualquier conjunto cerrado en $\mathbf{R}^p$ y considere a $H \cap K$; dado que este conjunto es cerrado y acotado (por el teorema 9.6*(c)*), el teorema de Heine-Borel asegura que $H \cap K$ es un subconjunto compacto de $\mathbf{R}^p$. Por el teorema 22.5 se concluye que $H_1 = f(H \cap K)$ es compacto y por lo tanto cerrado en $\mathbf{R}^q$. Si $g = f^{-1}$, entonces

$$H_1 = f(H \cap K) = g^{-1}(H).$$

Dado que H_1 es un subconjunto de $f(K) = D(g)$, esta última ecuación se puede escribir de la siguiente manera

$$H_1 \cap D(g) = g^{-1}(H).$$

Del teorema de continuidad global 22.1*(c)* se deduce que $g = f^{-1}$ es continua.

Q.E.D.

Se concluye esta sección introduciendo cierta notación que será necesaria.

22.10 DEFINICION. Si $D \subseteq \mathbf{R}^p$, entonces la colección de todas las funciones continuas de D a $\mathbf{R}^q$ se denota por medio de $C_{pq}(D)$. La colección de todas las **funciones continuas acotadas** de D a $\mathbf{R}^q$ se designa por medio de $BC_{pq}(D)$. Cuando p y q están sobreentendidos, estas colecciones se designarán simplemente como $C(D)$ y $BC(D)$.

La primera parte del siguiente resultado es una consecuencia del teorema 20.6 y la segunda parte se demuestra de igual manera que el lema 17.8.

22.11 TEOREMA. *(a) Los espacios $C_{pq}(D)$ y $BC_{pq}(D)$ son espacios vectoriales bajo las operaciones vectoriales*

$$(f+g)(x) = f(x) + g(x), \qquad (cf)(x) = cf(x) \qquad \text{for } x \in D.$$

(a) El espacio $BC_{pq}(D)$ es un espacio normado bajo la norma

$$\|f\|_D = \sup\{\|f(x)\| : x \in D\}.$$

Por supuesto, én el caso especial en que D es un subconjunto compacto de $\mathbf{R}^p$, $C_{pq}(D) = BC_{pq}(D)$.

Ejercicios

22.A. Interpretar el teorema de continuidad global 22.1 para las funciones de reloj real $f(x) = x^2$ y $g(x) = 1/x$, $x \neq 0$. Tomar varios conjuntos

abiertos y cerrados y determinar sus imágenes inversas bajo f y g.

22.B. Defínase $H: \boldsymbol{R} \to \boldsymbol{R}$ como

$$h(x) = 1, \qquad 0 \leq x \leq 1,$$
$$= 0, \qquad \text{en los otros casos.}$$

Dar un conjunto abierto G tal que $h^{-1}(G)$ no sea abierto en $\boldsymbol{R}$, y un conjunto cerrado F tal que $h^{-1}(F)$ no sea cerrado en $\boldsymbol{R}$.

22.C. Si f es acotada y continua de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ y si $f(x_0) > 0$, demostrar que f es estrictamente positiva en alguna vecindad de x_0. ¿Es válida la misma conclusión o para el caso en que f es simplemente continua en x_0?

22.D. Si $p: \boldsymbol{R}^2 \to \boldsymbol{R}$ es un polinomio y $c \in \boldsymbol{R}$, demostrar que el conjunto $\{(x, y): p(x, y) < c\}$ es abierto en $\boldsymbol{R}^2$.

22.E. Si $f: \boldsymbol{R}^p \to \boldsymbol{R}$ es continua en $\boldsymbol{R}^p$ y $\alpha < \beta$, demostrar que el conjunto $\{x \in \boldsymbol{R}^p: \alpha \leq f(x) \leq \beta\}$ es cerrado en $\boldsymbol{R}^p$.

22.F. Un subconjunto $D \subseteq \boldsymbol{R}^p$ es inconexo si y sólo si existe una función continua $f: D \to \boldsymbol{R}$ tal que $f(D) = \{0, 1\}$.

22.G. Sea f continua de $\boldsymbol{R}^2$ a $\boldsymbol{R}^q$. Defínanse las funciones g_1, g_2 de $\boldsymbol{R}$ a $\boldsymbol{R}^q$ como

$$g_1(t) = f(t, 0), \qquad g_2(t) = f(0, t).$$

Demostrar que g_1 y g_2 son continuas.

22.H. Relacione f, g_1, g_2 por medio de las fórmulas del ejercicio anterior. Demostrar que no se puede probar la continuidad de f en (0,0) a partir de la continuidad de g_1 y g_2 en $t = 0$

22.I. Dar un ejemplo de una función de $\boldsymbol{I} = [0, 1]$ a $\boldsymbol{R}$ que no sea acotada.

22.J. Dar un ejemplo de una función acotada f de $\boldsymbol{I}$ a $\boldsymbol{R}$ que no tome ninguno de los valores sup $\{f(x): x \in \boldsymbol{I}\}$ o inf $\{f(x): x \in \boldsymbol{I}\}$.

22.K. Dar un ejemplo de una función acotada y continua g de $\boldsymbol{R}$ a $\boldsymbol{R}$ que no tome ninguno de los valores sup $\{g(x): x \in \boldsymbol{R}\}$ o inf $\{g(x): x \in \boldsymbol{R}\}$.

22.L. Demostrar que todo polinomio de grado impar y coeficientes reales tiene una raíz real. Demostrar que el polinomio $p(x) = x^4 + 7x^3 - 9$ tiene cuando menos dos raíces reales.

22.M. Si $c > 0$ y n es un número natural, existe un único número positivo b tal que $b^n = c$.

22.N. Sea f continua de $\boldsymbol{I}$ a $\boldsymbol{R}$ con $f(0) < 0$ y $f(1) > 0$. Si $N = \{x \in \boldsymbol{I}: f(x) < 0\}$ y si $c = \sup N$, demostrar que $f(c) = 0$.

22.O. Sea f una función continua de $\boldsymbol{R}$ a $\boldsymbol{R}$ estrictamente creciente (en el sentido de que si $x' < x''$ entonces $f(x') < f(x'')$). Demostrar que f es inyectiva y que su función inversa f^{-1} es continua y estrictamente creciente.

22.P. Sea f una función continua de $\boldsymbol{R}$ a $\boldsymbol{R}$ que no toma ninguno de sus valores dos veces. ¿Es verdad que f debe ser estrictamente creciente o bien estrictamente decreciente?

22.Q. Sea g una función de $\boldsymbol{I}$ a $\boldsymbol{R}$. Demostrar que si g toma cada uno de sus valores exactamente dos veces, entonces g no puede ser continua en todo punto de $\boldsymbol{I}$.

22.R. Sea f continua en el intervalo $[0, 2\pi]$ a $\boldsymbol{R}$ y tal que $f(0) = f(2\pi)$. Demostrar que existe un punto c en este intervalo tal que $f(c) = f(c + \pi)$. (*Sugerencia:* considere $g(x) = f(x) - f(x + \pi)$.) Deducir que, en cualquier momento, existen puntos opuestos en el ecuador de la tierra que tienen la misma temperatura.

22.S. Defina $\varphi : [0, 2\pi) \to \boldsymbol{R}^2$ como $\varphi(t) = (\cos t, \sin t)$ para $t \in [0, 2\pi)$. Entonces φ es una aplicación continua inyectiva de $[0, 2\pi)$ sobre el círculo unitario $S = \{(x, y) \in \boldsymbol{R}^2 : x^2 + y^2 = 1\}$. Demostrar que $\varphi^{-1} : S \to [0, 2\pi)$ no no ser válido si el dominio no es compacto.)

Proyecto

22.xx El propósito de este proyecto es demostrar que muchos de los resultados de la sección 22 son válidos para funciones continuas cuyos dominios y rangos están contenidos en espacios métricos. (Al probar estos resultados se puede observar que definiciones anteriores son aplicables a espacios métricos o bien que se pueden reformular para que lo sean.)

(a) Probar que el teorema 20.2 se puede reformular para una función de un espacio métrico a otro.

(b) Demostrar que el teorema de continuidad global 22.1 es válido sin ningún cambio.

(c) Demostrar que es válido el teorema de conservación de conexidad 22.3

(d) Demostrar que es válido el teorema de conservación de compacidad 22.5

Sección 23 Continuidad uniforme y puntos fijos

Defínase f en un subconjunto $D(f)$ de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$. Es fácil ver que los siguientes resultados son equivalentes:

(i) f es continua en todo punto en $D(f)$.

(ii) Dada $\varepsilon > 0$ y $u \in D(f)$, existe una $\delta(\varepsilon, u) > 0$ tal que si x pertenece a $D(f)$ y $\|x - u\| \leq \delta$, entonces $\|f(x) - f(u)\| \leq \varepsilon$.

Lo que se debe observar aquí es que la δ depende, en general, tanto de ε como de u. El que δ dependa de u es consecuencia del hecho de que la función f puede cambiar sus valores con rapidez cerca de ciertos puntos y lentamente cerca de otros.

Puede suceder que la función sea tal que el número δ se pueda escoger independiente del punto u en $D(f)$ y dependiente únicamente de ε. Por ejemplo, si $f(x) = 2x$, entonces

$$|f(x) - f(u)| = 2\,|x - u|$$

de manera que se puede elegir $\delta(\varepsilon, u) = \varepsilon/2$ para todos los valores de u.

Por otro lado, si $g(x) = 1/x$ para $x > 0$, entonces

$$g(x) - g(u) = \frac{u - x}{ux}.$$

Si $0 < \delta < u$ y $|x - u| \leq \delta$, se deja al lector demostrar que

$$|g(x) - g(u)| \leq \frac{\delta}{u(u - \delta)}$$

y que esta desigualdad no se puede mejorar, ya que, de hecho, la igualdad es válida para $x = u - \delta$. Si se quiere hacer $|g(x) - g(u)| \leq \epsilon$, entonces el valor más grande para δ que se puede elegir es

$$\delta(\varepsilon, u) = \frac{\varepsilon u^2}{1 + \varepsilon u}.$$

De modo que si $u > 0$, g es continua en u ya que se puede elegir $\delta(\varepsilon, u) = \varepsilon u^2/(1 + \varepsilon u)$, y éste es el valor más grande que se puede escoger. Dado que

$$\inf\left\{\frac{\varepsilon u^2}{1 + \varepsilon u} : u > 0\right\} = 0,$$

no se puede obtener una $\delta(\varepsilon, u) > 0$ que sea independiente de la elección de u para todos los puntos $u > 0$.

Ahora se restringirá g a un dominio menor. De hecho, sea $a > 0$ y defínase $h(x) = 1/x$ para $x \geq a$. El análisis recién hecho prueba que se puede usar el mismo valor de $\delta(\varepsilon, u)$. Sin embargo, esta vez el dominio es más pequeño y

$$\inf\left\{\frac{\varepsilon u^2}{1 + \varepsilon u} : u \geq a\right\} = \frac{\varepsilon a^2}{1 + \varepsilon a} > 0.$$

De modo que si se define $\delta(\varepsilon) = \varepsilon a^2/(1 + \varepsilon a)$, se puede usar este número para todos los puntos $u \geq a$.

Para poder ordenar mejor estas ideas, el lector deberá revisar los ejemplos 20.5 y determinar en cuáles ejemplos la δ se eligió para que dependiera del punto en cuestión y en cuáles se eligió independientemente del punto.

Con esta introducción se establece ahora la definición formal.

23.1 DEFINICION. Sea f la función con dominio $D(f)$ en $\mathbf{R}^p$ y rango en $\mathbf{R}^q$. Se dice que f es **uniformemente continua** en un conjunto $A \subseteq D(f)$ si para cada $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si x y u pertenecen a A y $\|x - u\| \leq \delta(\varepsilon)$, entonces $\|f(x) - f(u)\| \leq \varepsilon$.

Es claro que si f es uniformemente continua en A entonces es continua en todo punto de A. Sin embargo, en general lo inverso no es válido. Es con-

veniente tener en mente qué se entiende al decir que una función no es uniformemente continua. Una vez establecido el siguiente criterio se deja al lector demostrarlo.

23.2 LEMA. *Una condición necesaria y suficiente para que la función f no sea uniformemente continua en $A \subseteq D(f)$ es que existan $\varepsilon_0 > 0$ y dos sucesiones $X = (x_n)$, $Y = (y_n)$ en A tales que si $n \in \mathbf{N}$, entonces $\|x_n - y_n\| \leq 1/n$ y $\|f(x_n) - f(y_n)\| > \varepsilon_0$.*

A manera de ejercicio, el lector deberá aplicar este criterio para demostrar que $g(x) = 1/x$ no es uniformemente continua en $D(g) = \{x : x > 0\}$.

En seguida se ofrece un resultado muy útil que asegura que una función continua de manera automática es uniformemente continua en cualquier conjunto compacto en su dominio.

23.3. TEOREMA DE CONTINUIDAD UNIFORME. *Sea f una función continua con dominio $D(f)$ en $\mathbf{R}^p$ y rango en $\mathbf{R}^q$. Si $K \subseteq D(f)$ es compacto, entonces f es uniformemente continua en K.*

PRIMERA DEMOSTRACION. Suponga que f no es uniformemente continua en K. Por el lema 23.2, existen $\varepsilon_0 > 0$ y dos sucesiones (x_n) y (y_n) en K tales que si $n \in \mathbf{N}$, entonces

$$\|x_n - y_n\| \leq 1/n, \qquad \|f(x_n) - f(y_n)\| > \varepsilon_0. \tag{23.1}$$

Como K es compacto en $\mathbf{R}^p$, la sucesión X es acotada; por el teorema de Bolzano-Weierstrass 16.4, hay una subsucesión $(x_{n(k)})$ de (x_n) que converge a un elemento z. Dado que K es cerrado, el límite z pertenece a K y f es continua en z. Es claro que la subsucesión correspondiente $(y_{n(k)})$ de Y también converge a z.

Del teorema 20.2*(c)* se deduce que ambas sucesiones $(f(x_{n(k)}))$ y $(f(y_{n(k)}))$ convergen a $f(z)$. Por lo tanto, cuando k es suficientemente grande, se tiene $\|f(x_{n(k)}) - f(y_{n(k)})\| < \varepsilon_0$. Pero esto contradice la segunda relación en (23.1).

SEGUNDA DEMOSTRACION. (Una demostración más corta se podría basar en el teorema de cobertura de Lebesgue 11.5; sin embargo, se ha convenido en usar la definición de compacidad) Supóngase que f es continua en todo punto del conjunto compacto K. Según el teorema 20.2*(b)*, dada $\varepsilon > 0$ y u en K hay un número $\delta(\frac{1}{2}\varepsilon, u) > 0$ tal que $x \in K$ y $\|x - u\| < \delta(\frac{1}{2}\varepsilon, u)$, entonces $\|f(x) - f(u)\| < \frac{1}{2}\varepsilon$. Para cada u en K, forme la bola abierta $G(u) = \{x \in \mathbf{R}^p : \|x - u\| < \frac{1}{2}\delta(\frac{1}{2}\varepsilon, u)\}$; entonces el conjunto K ciertamente está contenido en la unión de la familia $\mathcal{G} = \{G(u) : u \in K\}$, ya que para cada punto u en K hay una bola abierta $G(u)$ que lo contiene. Dado que K es compacto, está contenido en la unión de un número finito de conjuntos en la familia $\mathcal{G}$, digamos $G(u_1), \ldots, G(u_N)$. Se define ahora

$$\delta(\varepsilon) = \tfrac{1}{2} \inf \{\delta(\tfrac{1}{2}\varepsilon, u_1), \ldots, \delta(\tfrac{1}{2}\varepsilon, u_N)\},$$

y se habrá de probar que $\delta(\varepsilon)$ tiene la propiedad deseada. Suponga que x,u pertenecen a K y que $\|x-u\| \le \delta(\varepsilon)$ entonces existe un número natural k con $1 \le k \le N$ tal que x pertenece al conjunto $G(u_k)$; es decir, $\|x-u_k\| < \frac{1}{2}\delta(\frac{1}{2}\varepsilon, u_k)$. Dado que $\delta(\varepsilon) \le \frac{1}{2}\delta(\frac{1}{2}\varepsilon, u_k)$, se deduce que

$$\|u-u_k\| \le \|u-x\| + \|x-u_k\| < \delta(\tfrac{1}{2}\varepsilon, u_k).$$

Por lo tanto, se tienen las relaciones

$$\|f(x)-f(u_k)\| < \tfrac{1}{2}\varepsilon, \qquad \|f(u)-f(u_k)\| < \tfrac{1}{2}\varepsilon,$$

de donde $\|f(x)-f(u)\| < \varepsilon$. Se ha demostrado que si x,u son dos puntos arbitrarios de K para los cuales $\|x-u\| \le \delta(\varepsilon)$, entonces $\|f(x)-f(u)\| < \varepsilon$.

Q.E.D.

En secciones posteriores se hará uso en muchos casos del concepto de continuidad uniforme, de modo que en este momento no se darán aplicaciones. Sin embargo se introducirá otra propiedad que a menudo es asequible y suficiente para garantizar la continuidad uniforme.

23.4 DEFINICION. Si f tiene dominio $D(f)$ contenido en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$, se dice que f satisface una **condición de Lipschitz** † si existe una constante $A > 0$ tal que (23.2)

$$\|f(x)-f(u)\| \le A\,\|x-u\| \tag{23.2}$$

para todos los puntos x, u en $D(f)$. En el caso en que la desigualdad (23.2) sea válida para una constante $A < 1$, a la función de la llama una **contracción.**

Es claro que si la relación (23.2) es válida entonces al fijar $\delta(\varepsilon) = \varepsilon/A$ se puede establecer la continuidad uniforme de f en $D(f)$. Por lo tanto, si f satisface una condición de Lipschitz f es uniformemente continua. Sin embargo, lo inverso, no se cumple, como se puede ver al considerar la función definida para $D(f) = \boldsymbol{I}$ por medio de $f(x) = \sqrt{x}$. Si se cumple (23.2), entonces, fijando $u = 0$ se debe tener $|f(x)| \le A\,|x|$ para alguna constante A, pero fácilmente se ve que la última desigualdad no se puede satisfacer.

Recordando el teorema 21.3 se puede ver que una función lineal con dominio $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$ satisface una condición de Lipschitz. Más aún, en la sección 27 se verá que cualquier función real con derivada acotada también satisface una condición de Lipschitz.

Teoremas del punto fijo

Si f es una función con dominio $D(f)$ y rango en el mismo espacio $\boldsymbol{R}^p$, entonces se dice que un punto u en $D(f)$ es un **punto fijo** de f siempre que

†RUDOLPH LIPSCHITZ (1832-1903) fue profesor en Bonn. Hizo aportaciones en álgebra, teoría de números, geometría diferencial y análisis.

$f(u) = u$. Muchos resultados importantes se pueden demostrar con base en la existencia de puntos fijos de funciones, de modo que resulta importante tener cierto criterio positivo en este sentido. El primer teorema que se da es de índole elemental; sin embargo, a menudo es útil y tiene la importante ventaja de que proporciona una construcción para el punto fijo. Para simplificar, primero se dará el resultado para el caso en que el dominio de la función es todo el espacio.

23.5 TEOREMA DEL PUNTO FIJO PARA CONTRACCIONES. *Sea f una contracción con dominio $\mathbf{R}^p$ y rango contenido en $\mathbf{R}^p$. Entonces, f tiene un punto fijo único.*

DEMOSTRACION. Se está suponiendo que existe una constante C con $0 < C < 1$ tal que $\|f(x) - f(y)\| \leq C\,\|x - y\|$ para todas x, y en $\mathbf{R}^p$. Sea x_1 un punto arbitrario en $\mathbf{R}^p$ y fije $x_2 = f(x_1)$; inductivamente, fije.

$$x_{n+1} = f(x_n), \qquad n \in \mathbf{N}. \tag{23.3}$$

Se habrá de probar que la sucesión (x_n) converge a un punto fijo único de f y se calculará la rapidez de la convergencia.

Para hacer esto, observe que

$$\|x_3 - x_2\| = \|f(x_2) - f(x_1)\| \leq C\,\|x_2 - x_1\|,$$

e, inductivamente, que

$$\|x_{n+1} - x_n\| = \|f(x_n) - f(x_{n-1})\| \leq C\,\|x_n - x_{n-1}\| \leq C^{n-1}\,\|x_2 - x_1\|. \tag{23.4}$$

Si $m \geq n$, el uso repetido de (23.4) de

$$\begin{aligned}\|x_m - x_n\| &\leq \|x_m - x_{m-1}\| + \|x_{m-1} - x_{m-2}\| + \cdots + \|x_{n+1} - x_n\| \\ &\leq \{C^{m-2} + C^{m-3} + \cdots + C^{n-1}\}\,\|x_2 - x_1\|.\end{aligned}$$

Por lo que se deduce que, para $m \geq n$,

$$\|x_m - x_n\| \leq \frac{C^{n-1}}{1 - C}\,\|x_2 - x_1\|. \tag{23.5}$$

Puesto que $0 < C < 1$, la sucesión (C^{n-1}) converge a cero. Por lo tanto, (x_n) es una sucesión de Cauchy. Si $u = \lim (x_n)$, de (23.3) es claro que u es un punto fijo de f. De (23.5) y el lema 15.8 se obtiene el cálculo

$$\|u - x_n\| \leq \frac{C^{n-1}}{1 - C}\,\|x_2 - x_1\| \tag{23.6}$$

para la rapidez de la convergencia.

Por último, se demuestra que sólo hay un punto fijo para f. De hecho, si u, v son dos puntos fijos, distintos, de f, entonces

$$\|u - v\| = \|f(u) - f(v)\| \leq C\,\|u - v\|.$$

Puesto que $u \neq v$, entonces $\|u - v\| \neq 0$, por lo que esta relación implica que $1 \leq C$, contrario a la hipótesis de que $C < 1$. Q.E.D.

Note que lo que realmenre se ha hecho es demostrar el siguiente resultado.

23.6 COROLARIO *Si f es una contracción con* $C < 1$, *constante si* x_1 *es un punto arbitrario en* $\mathbf{R}^p$, *y si la sucesión* $X = (x_n)$ *está definida por la ecuación (23.3), entonces X converge al punto fijo único u de f con la rapidez calculada en (23.6).*

Para el caso en que la función f no este definida en todo $\mathbf{R}^p$, se deberá tener más cuidado para asegurar que la definición iterativa (23.3) de la sucesión se puede llevar a cabo y que los puntos permanezcan en el dominio de f. Aun cuando son posibles otras formulaciones, habrá que conformarse con la siguiente.

23.7 TEOREMA. *Suponga que f es una contracción con C constante definida para* $D(f) = \{x \in \mathbf{R}^p : \|x\| \leq B\}$ *y que* $\|f(0)\| \leq B(1 - C)$. *Entonces la sucesión*

$$x_1 = 0, x_2 = f(x_1), \ldots, x_{n+1} = f(x_n), \ldots$$

converge al punto fijo único de f que se encuentra en el conjunto D(f).

DEMOSTRACION. En efecto, si $x \in D = D(f)$, entonces $\|f(x) - f(0)\| \leq C\,\|x - 0\| \leq CB$, por lo que se deduce que

$$\|f(x)\| \leq \|f(0)\| + CB \leq (1 - C)B + CB = B.$$

Por lo tanto, $f(D) \subseteq D$. De modo que la sucesión (x_n) se puede definir y permanece en D y la demostración anterior es aplicable. Q.E.D.

El teorema de contracción que se acaba de establecer tiene ciertas ventajas: es constructivo, el error de aproximación se puede calcular y garantiza un punto fijo único. Sin embargo, tiene la desventaja de que el requisito de que f sea una contracción es una restricción muy severa. Un hecho importane y profundo demostrado por primera vez en 1910 por L.E.J. Brouwer,† es que *cualquier* función continua con dominio $D = \{x \in \mathbf{R}^p : \|x\| \leq B\}$ y rango contenido en D debe tener cuando menos un punto fijo.

23.8 TEOREMA DEL PUNTO FIJO DE BROWER. *Sea* $B > 0$ *y sea* $D = \{x \in \mathbf{R}^p : \|x\| \leq B\}$. *Entonces cualquier función continua con dominio D y rango contenido en D tiene cuando menos un punto fijo.*

La demostración de este resultado para $p = 1$ se dará como ejercicio. Para el caso $p > 1$, no obstante, la demostración se alejaría mucho del tema.

†L. E. J. BROUWER (1881-1966) fue profesor en Amsterdam y decano de la escuela holandesa de matemáticas. Además de sus aportaciones a la topología, se le conoce por su trabajo en los fundamentos de las matemáticas.

Para una demostración basada sólo en conceptos elementales, véase Dunford-Schwartz, págs. 467-470. Para un informe más sistematizado de punto fijo y teoremas relacionados, consulte el libro de Lifschitz.

Ejercicios

23.A. Analizar cada una de las funciones del ejemplo 20.5 y probar que la función es uniformemente continua en su dominio o bien que no lo es.

23.B. Dar una demostración del teorema de continuidad uniforme 23.3 usando el teorema de cobertura de Lebesgue 11.5.

23.C. Si B está acotado en $\boldsymbol{R}^p$ y $f: B \to \boldsymbol{R}^q$ es uniformemente continua, demostrar que f está acotada en B. Demostrar que esta conclusión no es válida si B no es acotado en $\boldsymbol{R}^p$.

23.D. Demostrar que las funciones definidas para $x \in \boldsymbol{R}$ por medio de

$$f(x) = \frac{1}{1+x^2}, \qquad g(x) = \sin x,$$

son uniformemente continuas en $\boldsymbol{R}$.

23.E. Demostrar que las funciones definidas para $D = \{x \in \boldsymbol{R} : x \geq 0\}$, por medio de

$$h(x) = x, \qquad k(x) = e^{-x},$$

son uniformemente continuas en D.

23.F. Demostrar que las siguientes funciones no son uniformemente continuas en sus dominios.

(a) $f(x) = 1/x^2$, $\quad D(f) = \{x \in \boldsymbol{R} : x > 0\}$,
(b) $g(x) = \tan x$, $\quad D(g) = \{x \in \boldsymbol{R} : 0 \leq x < \pi/2\}$,
(c) $h(x) = e^x$, $\quad D(h) = \boldsymbol{R}$.
(d) $k(x) = \sin(1/x)$, $\quad D(k) = \{x \in \boldsymbol{R} : x > 0\}$.

23.G. Una función $g: \boldsymbol{R} \to \boldsymbol{R}^q$ es periódica si existe un número $p > 0$ tal que $g(x+p) = g(x)$ para toda $x \in \boldsymbol{R}$. Demostrar que una función periódica continua es acotada y uniformemente continua en $\boldsymbol{R}$.

23.H. Defínase f de $D \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^q$, y suponga que f es uniformemente continua en D. Si (x_n) es una sucesión de Cauchy en D, demostrar que $(f(x_n))$ es una sucesión de Cauchy en $\boldsymbol{R}^q$.

23.I. Suponga que $f: (0, 1) \to \boldsymbol{R}$ es uniformemente continua en (0,1). Demostrar que f se puede definir en $x = 0$ y $x = 1$ de tal manera que se vuelve continua en [0, 1].

23.J. Sea $D = \{x \in \boldsymbol{R}^p : \|x\| < 1\}$. Demostrar que $f: D \to \boldsymbol{R}^q$ se puede extender a una función continua de $D_1 = \{x \in \boldsymbol{R}^p : \|x\| \leq 1\}$ a $\boldsymbol{R}^q$ si y sólo si f es uniformemente continua en D.

23.K. Si f y g son uniformemente continuas de $\boldsymbol{R}$ a $\boldsymbol{R}$, demostrar que g es uniformemente continua en $\boldsymbol{R}$, pero fg no necesariamente es uniformemente continua en $\boldsymbol{R}$ aun cuando alguna de las dos, f o g, esté acotada.

23.L. Si $f: \boldsymbol{I} \to \boldsymbol{I}$ es continua, demostrar que f tiene en punto fijo en I. (*Sugerencia:* Considerar $g(x) = f(x) - x$.)

23.M. Dar un ejemplo de una función $f: \boldsymbol{R}^p \to \boldsymbol{R}^p$ tal que $\|f(x) - f(u)\| \leq \|x - u\|$ para todas $x, u \in \boldsymbol{R}^p$ que no tenga un punto fijo. (¿por qué no contradice esto al teorema 23.5?)

23.N. Sean f y g funciones continuas $[a, b]$ tales que el rango $R(f) \subseteq R(g) = [0, 1]$. Demostrar que existe un punto $c \in [a, b]$ tal que $f(c) = g(c)$.

Proyecto

23.α. Este proyecto introduce el concepto de la "oscilación" de una función en un conjunto y en un punto. Sea $I = [a, b] \subseteq \mathbf{R}$ y sea $f : I \to \mathbf{R}$ acotada. Si $A \subseteq I$ se define la oscilación de f en A como el número

$$\Omega_f(A) = \sup \{f(x) - f(y) : x, y \in A\}.$$

(a) Demostrar que $0 \leq \Omega_f(A) \leq 2 \sup \{|f(x)| : x \in A\}$. Si $A \subseteq B \subseteq I$, entonces $\Omega_f(A) \leq \Omega_f(B)$.

(b) Si $c \in I$, la *oscilación de f en c* se define por medio del número

$$\omega_f(c) = \inf_{\delta} \Omega_f(N_\delta)$$

en donde $N_\delta = \{x \in I : |x - x_0| < \delta\}$ demostrar que

$$\omega_f(c) = \lim_{\delta \to 0} \Omega_f(N_\delta).$$

Además, si $\omega_f(c) < \alpha$, entonces existe $\delta > 0$ tal que $\Omega_f(N_\delta) < \alpha$.

(c) Demostrar que f es continua en $c \in I$ si y sólo si $\omega_f(c) = 0$.

(d) Si $\alpha > 0$ y si $\omega_f(x) < \alpha$ para toda $x \in I$, entonces existe $\delta > 0$ tal que su diámetro $d(A) = \sup \{|x - y| : x, y \in A\}$ es menor que δ, entonces $\Omega_f(A) < \alpha$.

(e) Si $\alpha > 0$, entonces el conjunto $D_\alpha = \{x \in I : \omega_f(x) \geq \alpha\}$ es un conjunto cerrado en $\mathbf{R}$. Demostrar que

$$D = \bigcup_{\alpha > 0} D_\alpha = \bigcup_{n \in N} D_{1/n}$$

es el conjunto de puntos en donde f es discontinua. De donde el conjunto de puntos de discontinuidad de una función es la unión de una familia contable de conjuntos cerrados. (Dicho conjunto recibe el nombre de conjunto F_σ.)

(f) Ampliar estas definiciones y resultados para una función definida en una celda cerrada en $\mathbf{R}^p$.

Sección 24 Sucesiones de funciones continuas

En muchas ocasiones es necesario considerar una sucesión de funciones continuas. En esta sección se dan varios teoremas interesantes e importantes acerca de dichas sucesiones. El teorema 24.1 es un resultado clave que se usará a menudo en lo sucesivo. Los demás teoremas no se usarán con frecuencia; sin embargo, el lector deberá, cuando menos, familiarizarse con los enunciados.

En esta sección se deberá aclarar la importancia de la convergencia uniforme. Recuerde que una sucesión (f_n) de funciones en un subconjunto D de

$\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ se dice que converge uniformemente en D a f si para toda $\varepsilon > 0$ existe una $N(\varepsilon)$ tal que si $n \geq N(\varepsilon)$ y $x \in D$, entonces $\|f_n(x) - f(x)\| < \varepsilon$. De nuevo, por el teorema 17.9, esto es válido si y sólo si $\|f_n - f\|_D \to 0$, cuando (f_n) es una sucesión acotada.

Intercambio de límite y continuidad

Observe que el límite de una sucesión de funciones continuas puede no ser continuo. Es muy fácil ver esto; para $n \in \boldsymbol{N}$, y $x \in \boldsymbol{I}$, $f_n(x) = x^n$. En el ejemplo 17.2(b) se vio que la sucesión (f_n) converge en $\boldsymbol{I}$ a la función f definida por

$$\begin{aligned} f(x) &= 0, \qquad 0 \leq x < 1, \\ &= 1, \qquad x = 1. \end{aligned}$$

Por lo que, a pesar de la índole tan sencilla de las funciones continuas f_n, la función límite no es continua en el punto $x = 1$.

Aun cuando la extensión de la discontinuidad de la función límite en el ejemplo recién dado no es muy grande, es evidente que se pueden construir ejemplos más complicados en los que se obtenga una discontinuidad más extensa. Sería interesante investigar qué tan discontinuo puede ser el límite de una sucesión de funciones continuas; sin embargo, esta investigación se saldría mucho del campo. Más aún, para la mayoría de las aplicaciones es más importante encontrar condiciones adicionales que garanticen que la función límite es continua.

En seguida se probará el hecho de que la convergencia uniforme de una sucesión de funciones continuas es suficiente para garantizar la continuidad de la función límite.

24.1 TEOREMA. *Sea $F = (f_n)$ una sucesión de funciones continuas con dominio D en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$ tal que converge uniformemente en D a la función f. Entonces f es continua en D.*

DEMOSTRACION. Como (f_n) converge uniformemente en D a f, dada $\varepsilon > 0$ existe un número natural $N = N(\varepsilon/3)$ tal que $\|f_N(x) - f(x)\| < \varepsilon/3$ para toda x en D. Para probar que f es continua en un punto a en D, observe que

$$(24.1) \qquad \begin{aligned} \|f(x) - f(a)\| &\leq \|f(x) - f_N(x)\| + \|f_N(x) - f_N(a)\| + \|f_N(a) - f(a)\| \\ &\leq \varepsilon/3 + \|f_N(x) - f_N(a)\| + \varepsilon/3. \end{aligned}$$

Dado que f_N es continua, existe un número $\delta = \delta(\varepsilon/3, a, f_N) > 0$ tal que si $\|x - a\| < \delta$ y $x \in D$, entonces $\|f_N(x) - f_N(a)\| < \varepsilon/3$. (Véase la figura 24.1.) Por lo tanto, para tal x se tiene $\|f(x) - f(a)\| < \varepsilon$. Esto prueba la continuidad de la función f en un punto arbitrario a en D. Q.E.D.

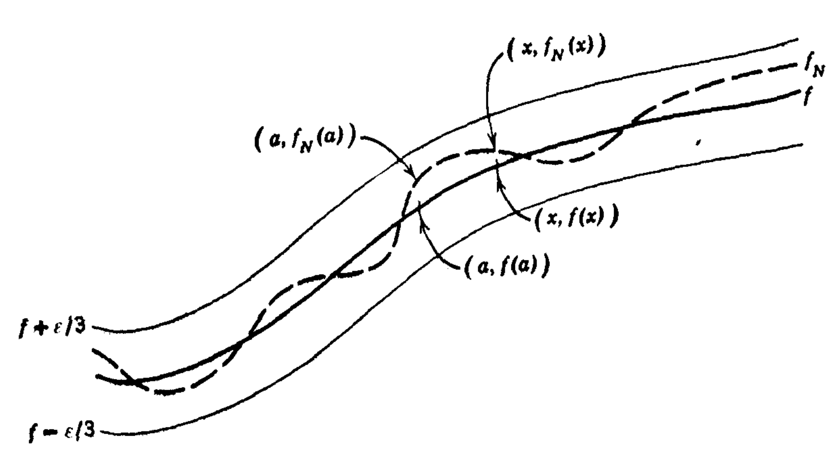


Figura 24.1

Se hace la aclaración de que aun cuando la convergencia uniforme de la sucesión de funciones continuas es suficiente para la continuidad de la función límite, no es necesaria. De modo que (f_n) es una sucesión de funciones continuas que converge a una función continua f, no se deduce que la convergencia sea uniforme (véase el ejercicio 24.A).

Como se vio en el teorema 17.9, la convergencia uniforme en un conjunto D de una sucesión de funciones está implícito por la convergencia en la norma uniforme en D. Por lo tanto, el teorema 24.1 se formula de la siguiente manera.

24.2 TEOREMA. Si (f_n) es una sucesión de funciones en $BC_{pq}(D)$ tal que $\|f_n - f\|_D \to 0$, entonces $f \in BC_{pq}(D)$.

Teoremas de aproximación

En muchas aplicaciones es conveniente "aproximar" funciones continuas por medio de funciones de naturaleza elemental. A pesar de que existen varias definiciones razonables que se pueden emplear para hacer más precisa la palabra "aproximar", una de las más naturales, así como importante, es la que exige que en todo punto del dominio dado la función de aproximación no debe diferir de la función dada en más del error prefijado. A este juicio algunas veces se le conoce como "aproximación uniforme" y está íntimamente ligado a la convergencia uniforme. Suponiendo que f es una función dada con dominio $D = D(f)$, contenido en $\mathbf{R}^p$ y con rango en $\mathbf{R}^q$, se dice que una función g **aproxima a** f **uniformemente en** D **en** $\varepsilon > 0$, si

$$\|g(x) - f(x)\| \le \varepsilon \qquad \text{para toda } x \in D;$$

o de modo equivalente si

$$\|g - f\|_D = \sup \{\|g(x) - f(x)\| : x \in D\} \le \varepsilon.$$

Se ha usado en este caso la norma que se introdujo en la ecuación (17.5). Se dice que la función f se puede **aproximar uniformemente en** D por funciones en una clase $\mathcal{G}$ si para cada número $\varepsilon > 0$ hay una función g_ε en $\mathcal{G}$ tal que $\|g_\varepsilon - f\|_D < \varepsilon$; o, lo que es lo mismo, si existe una sucesión de funciones en $\mathcal{G}$ que converge uniformemente en D a f.

24.3 DEFINICION. A una función g con dominio $\mathbf{R}^p$ y rango en $\mathbf{R}^q$ se le llama una **función escalonada** si toma únicamente un número finito de valores distintos en $\mathbf{R}^q$, tomándose cada valor distinto de cero en un intervalo en $\mathbf{R}^p$.

Por ejemplo, si $p = q = 1$, entonces la función g definida explícitamente por

$$\begin{aligned} g(x) &= 0, && x \leq -2, \\ &= 1, && -2 < x \leq 0, \\ &= 3, && 0 < x < 1, \\ &= -5, && 1 \leq x \leq 3, \\ &= 0, && x > 3. \end{aligned}$$

es una función escalonada.

En seguida se demuestra que una función continua cuyo dominio es una celda compacta se puede aproximar uniformemente por medio de funciones escalonadas.

24.4 TEOREMA. *Sea f una función continua cuyo dominio D es una celda compacta en $\mathbf{R}^p$ y cuyos valores pertenecen a $\mathbf{R}^q$. Entonces f se puede aproximar uniformemente en D por medio de funciones escalonadas.*

DEMOSTRACION. Sea $\varepsilon > 0$ dada; como f es uniformemente continua (teorema 23.3), existe un número $\delta(\varepsilon) > 0$ tal que si x, y pertenecen a $\|x - y\| < \delta(\varepsilon)$, entonces $\|f(x) - f(y)\| < \varepsilon$. Divida el dominio D de f en celdas ajenas $I_1, \ldots, I_n$ tales que si x, y pertenecen a I_k, entonces $\|x - y\| < \delta(\varepsilon)$. (¿cómo?) Sea x_k cualquier punto perteneciente a la celda I_k, $k = 1, \ldots, n$ y defínanse $g_\varepsilon(x) = f(x_k)$ para $x \in I_k$ y $g_\varepsilon(x) = 0$ para $x \notin D$. Es claro entonces que $\|g_\varepsilon(x) - f(x)\| < \varepsilon$ para $x \in D$ de tal manera que g_ε aproxima a f uniformemente en D en no más de ε. (figura 24.2.) Q.E.D.

Es natural suponer que una función continua se puede aproximar uniformemente por funciones simples que también sean continuas (las funciones escalonadas no lo son). Para hacerlo más sencillo, el siguiente resultado se dará sólo para el caso en que $p = q = 1$ aunque es evidente que hay una generalización para dimensiones más altas.

Se dice que una función g definida en una celda compacta $J = [a, b]$ de $\mathbf{R}$ con valores en $\mathbf{R}$ es **lineal por partes** si hay un número finito de puntos c_k

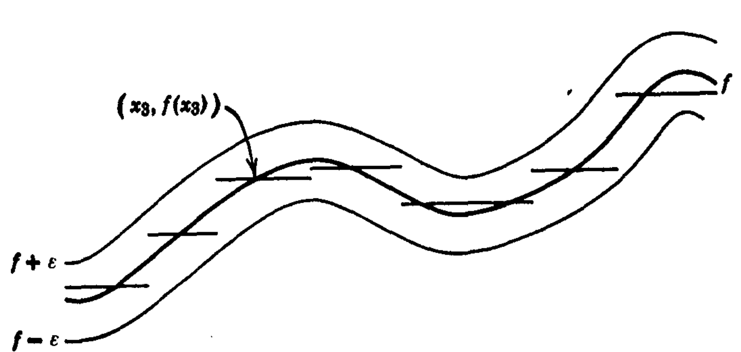


Figura. 24.2. Aproximación por medio de una función escalonada.

con $a = c_0 < c_1 < c_2 < \cdots < c_n = b$ y números reales correspondientes $A_k, B_k, k = 0, 1, \ldots, n$, tales que cuando x satisface la relación $c_{k-1} < x < c_k$, la función g es de la forma

$$g(x) = A_k x + B_k, \qquad k = 0, 1, \ldots, n.$$

Si g es continua en J, las constantes A_k, B_k deben satisfacer, desde luego, ciertas relaciones.

24.5 TEOREMA. Sea f una función continua cuyo dominio es una celda compacta J en $\boldsymbol{R}$. Entonces f se puede aproximar uniformemente en J por medio de funciones continuas lineales por partes.

DEMOSTRACION. Como en el caso anterior, f es uniformemente continua en el conjunto compacto J. Por lo tanto, dada $\varepsilon > 0$, se divide $J = [a, b]$ en celdas agregando puntos intermedios $c_k, k = 0, 1, \ldots, n$, con $a = c_0 < c_1 < c_2 < \cdots < c_n = b$ de tal manera que $c_k - c_{k-1} < \delta(\varepsilon)$. Una los puntos $(c_k, f(c_k))$ mediante segmentos de línea y defina la función continua lineal por partes g que resulta. Es claro que g aproxima a f uniformemente en J en no más de ε. Q.E.D.

Aproximación por polinomios

Se demostrará ahora un resultado más profundo, útil e interesante que se refiere a la aproximación por polinomios. Primero se demuestra el teorema de aproximación de Weierstrass para $p = q = 1$, usando los polinomios de S. Bernstein.†

24.6 DEFINICION. Sea f una función con dominio $\boldsymbol{I} = [0, 1]$ y rango en $\boldsymbol{R}$. El enésimo polinomio de Bernstein para f se define como (24.2)

†SERGE N. BERNSTEIN (1880-1968) hizo grandes aportaciones al análisis, a la teoría de aproximación y a la probabilidad. Nació en Odessa y fue profesor en Leningrado y en Moscú.

(24.2) $$B_n(x) = B_n(x; f) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right)\binom{n}{k} x^k (1-x)^{n-k}.$$

Estos polinomios de Bernstein no son tan complicados como parecen serlo a primera vista. El lector que tenga cierta experiencia en probabilidad podrá ver que detrás de esto está el binomio de distribución. Aun sin este tipo de experiencia, el lector deberá observar que el valor $B_n(x; f)$ del polinomio en el punto x se calcula a partir de los valores $f(0)$, $f(1/n)$, $f(2/n), \ldots, f(1)$, con ciertos valores de peso no negativos $\varphi_k(x) = \binom{n}{k} x^k (1-x)^{n-k}$ que se podrán ver muy pequeños para aquellos valores de k para los cuales k/n está lejos de x. De hecho, la función φ_k es no negativa en I y toma su valor máximo en el punto k/n. Más aún, como se verá en seguida, la suma de todas las $\varphi_k(x)$, $k = 0, 1, \ldots, n$, es 1 para cada x en I.

Recuerde que el teorema del binomio asegura que (24.3)

(24.3) $$(s+t)^n = \sum_{k=0}^{n} \binom{n}{k} s^k t^{n-k},$$

en donde $\binom{n}{k}$ denota al coeficiente binomial

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Directamente se puede observar que

(24.4) $$\binom{n-1}{k-1} = \frac{(n-1)!}{(k-1)!(n-k)!} = \frac{k}{n}\binom{n}{k},$$

(24.5) $$\binom{n-2}{k-2} = \frac{(n-2)!}{(k-2)!(n-k)!} = \frac{k(k-1)}{n(n-1)}\binom{n}{k}.$$

Ahora, sean $s = x$ y $t = 1 - x$ en (24.3) para obtener

(24.6) $$1 = \sum_{k=0}^{n} \binom{n}{k} x^k (1-x)^{n-k}.$$

Escribiendo (24.6) con n reemplazada por $n-1$ y k por j, se tiene

$$1 = \sum_{j=0}^{n-1} \binom{n-1}{j} x^j (1-x)^{n-1-j}.$$

Multiplicar esta última relación por x y aplicar la igualdad (24.4) para obtener

$$x = \sum_{j=0}^{n-1} \frac{j+1}{n} \binom{n}{j+1} x^{j+1} (1-x)^{n-(j+1)}.$$

Ahora sea $k = j + 1$, entonces

$$x = \sum_{k=1}^{n} \frac{k}{n}\binom{n}{k} x^k (1-x)^{n-k}.$$

Observe también que el término correspondiente a $k=0$ se puede incluir, que es cero. Por lo tanto, se tiene

$$x = \sum_{k=0}^{n} \frac{k}{n}\binom{n}{k} x^k (1-x)^{n-k}. \tag{24.7}$$

Un cálculo análogo basado en (24.6) con n reemplazada por $n-2$ y la igualdad (24.5) prueba que

$$(n^2-n)x^2 = \sum_{k=0}^{n} (k^2-k)\binom{n}{k} x^k (1-x)^{n-k}.$$

Por lo tanto, se concluye que

$$\left(1-\frac{1}{n}\right)x^2 + \frac{1}{n}x = \sum_{k=0}^{n} \left(\frac{k}{n}\right)^2 \binom{n}{k} x^k (1-x)^{n-k}. \tag{24.8}$$

Multiplicando (24.6) por x^2, (24.7) por $-2x$, y sumándolas a (24.8) se obtiene

$$(1/n)x(1-x) = \sum_{k=0}^{n} (x-k/n)^2 \binom{n}{k} x^k (1-x)^{n-k}, \tag{24.9}$$

que es un cálculo que se necesitará más adelante.

Examinando la definición 24.6, la fórmula (24.6) dice que el enésimo polinomio de Bernstein para la función constante $f_0(x)=1$ coincide con f_0. La fórmula (24.7) dice lo mismo para la función $f_1(x)=x$. La fórmula (24.8) asegura que el enésimo polinomio de Bernstein para la función $f_2(x)=x^2$ es

$$B_n(x; f_2) = (1-1/n)x^2 + (1/n)x,$$

que converge uniformemente en $\boldsymbol{I}$ a f_2. En seguida se demostrará que si f es cualquier función continua de $\boldsymbol{I}$ a $\boldsymbol{R}$ entonces la sucesión de polinomios de Bernstein tiene la propiedad de que converge uniformemente en $\boldsymbol{I}$ a f. Esto nos dará una demostración constructiva del teorema de aproximación de Weierstrass. En el desarrollo de esta demostración será necesaria la fórmula (24 9).

24.7 TEOREMA DE APROXIMACION DE BERNSTEIN. *Sea f continua en $\boldsymbol{I}$ con valores en $\boldsymbol{R}$. Entonces la sucesión de polinomios de Bernstein para f definida en la ecuación (24.2) converge uniformemente en $\boldsymbol{I}$ a f.*

DEMOSTRACION. Al multiplicar la fórmula (24.6) por $f(x)$ se obtiene

$$f(x) = \sum_{k=0}^{n} f(x)\binom{n}{k} x^k (1-x)^{n-k}.$$

Por lo tanto, se obtiene la relación

$$f(x)-B_n(x)=\sum_{k=0}^{n}\{f(x)-f(k/n)\}\binom{n}{k}x^k(1-x)^{n-k}$$

de donde se deduce que

$$(24.10)\qquad |f(x)-B_n(x)|\leq\sum_{k=0}^{n}|f(x)-f(k/n)|\binom{n}{k}x^k(1-x)^{n-k}.$$

f está acotada, digamos que por M, y también es uniformemente continua. Observe que si k es tal que k/n está cerca de x entonces el término correspondiente en la suma (24.10) es pequeño por la continuidad de f en x; por otro lado, si k/n está lejos de x, el factor en relación con f sólo se puede decir que es menor que $2M$ y cualquier pequeñez debe surgir de los otros factores. Por lo tanto, el camino a seguir es dividir (24.10) en dos partes: aquellos valores de k en donde $x-k/n$ es pequeño y aquellos en los que $x-k/n$ es grande.

Sea $\varepsilon>0$ y sea $\delta(\varepsilon)$ como en la definición de continuidad uniforme de f. Resulta conveniente elegir n de tal magnitud que

$$(24.11)\qquad n\geq\sup\{(\delta(\varepsilon))^{-4},\ M^2/\varepsilon^2\},$$

y partir (24.10) en dos sumas. La suma tomada sobre aquellas k para las cuales $|x-k/n|<n^{-1/4}\leq\delta(\varepsilon)$ da

$$\sum_k\varepsilon\binom{n}{k}x^k(1-x)^{n-k}\leq\varepsilon\sum_{k=1}^{n}\binom{n}{k}x^k(1-x)^{n-k}=\varepsilon.$$

La suma tomada sobre aquellas k para las cuales $|x-k/n|\geq n^{-1/4}$, es decir $(x-k/m)^2\geq n^{-1/2}$, se puede calcular usando la fórmula (24.9). Para esta parte de la suma en (24.10) se obtiene la cota superior

$$\begin{aligned}\sum_k 2M\binom{n}{k}x^k(1-x)^{n-k}&=2M\sum_k\frac{(x-k/n)^2}{(x-k/n)^2}\binom{n}{k}x^k(1-x)^{n-k}\\&\leq 2M\sqrt{n}\sum_{k=1}^{n}(x-k/n)^2\binom{n}{k}x^k(1-x)^{n-k}\\&\leq 2M\sqrt{n}\left\{\frac{1}{n}x(1-x)\right\}\leq\frac{M}{2\sqrt{n}},\end{aligned}$$

ya que $x(1-x)\leq\frac{1}{4}$ en el intervalo I. Recordando la resolución (24.11) para n, se concluye que cada una de estas dos partes de (24.10) está acotada por arriba por ε. Por lo tanto, para alguna n elegida en (24.11) se tiene

$$|f(x)-B_n(x)|<2\varepsilon,$$

independientemente del valor de x. Esto prueba que la sucesión (B_n) converge uniformemente en I a f. Q.E.D.

Como corolario directo del teorema de Bernstein se tiene el siguiente resultado importante.

24.8 TEOREMA DE APROXIMACION DE WEIERSTRASS. *Sea f una función continua en un intervalo compacto de $\boldsymbol{R}$ y con valores en R. Entonces f se puede aproximar uniformemente por polinomios.*

DEMOSTRACION. Si f está definida en $[a, b]$, entonces la función g definida en $\boldsymbol{I} = [0, 1]$ por medio de

$$g(t) = f((b-a)t + a), \qquad t \in \boldsymbol{I},$$

es continua. Por lo tanto, g se puede aproximar uniformemente por polinomios de Bernstein y un simple cambio de variable da una aproximación polinomial de f. Q.E.D.

Se decidió repasar los detalles de la demostración del teorema de Bernstein 24.7 porque proporciona un método constructivo para encontrar una sucesión de polinomios que converja uniformemente en $\boldsymbol{I}$ a la función continua dada. Además, el método de la demostración del teorema 24.6 es típico de muchos argumentos analíticos y es importante desarrollar un entendimiento de dichos argumentos. Por último, aun cuando en la sección 26 se establecerán resultados más generales de aproximación, para poder hacerlo es necesario saber que la función valor absoluto se puede aproximar uniformemente en un intervalo compacto por polinomios. Aun cuando sería posible probar este caso especial en forma directa, la demostración no es tan sencilla. Para un análisis más completo de aproximación el lector podrá dirigirse al libro de E. Cheney que aparece en la lista de libros de consulta.

Ejercicios

24.A. Dar un ejemplo de una sucesión de funciones continuas que converja a una función continua pero en donde la convergencia no sea uniforme.

24.B. Dar un ejemplo de una sucesión de funciones discontinuas en todas partes que converja uniformemente a una función continua.

24.C. Dar un ejemplo de una sucesión de funciones continuas que converja, en un conjunto compacto, a una función que tenga un número infinito de discontinuidades.

24.D. Sea (f_n) una sucesión de funciones continuas de $D \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^q$ tal que (f_n) converge uniformemente a f en D y sea (x_n) una sucesión de elementos en D que converge a $x \in D$. ¿Se deduce que $(f_n(x_n))$ converge a $f(x)$?

24.E. Considere las sucesiones (f_n) definidas en $D = \{x \in \boldsymbol{R} : x \geq 0\}$ a $\boldsymbol{R}$ por medio de las siguientes fórmulas:

(a) $\dfrac{x^n}{n}$, (b) $\dfrac{x^n}{1+x^n}$, (c) $\dfrac{x^n}{n+x^n}$,

(d) $\dfrac{x^{2n}}{1+x^n}$, (e) $\dfrac{x^n}{1+x^{2n}}$, (f) $\dfrac{x}{n}e^{-x/n}$.

Analizar la convergencia y la convergencia uniforme de estas sucesiones y la continuidad de las funciones límite. En caso de que no haya convergencia uniforme en D, considere los intervalos apropiados en D.

24.F. Sea (f_n) una sucesión de $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ que converge en D a f. Suponga que cada f_n es continua en c y que la sucesión converge uniformemente en alguna vecindad de c. Demostrar que f es continua en c.

24.G. Sea (f_n) una sucesión de funciones continuas de $D \subseteq \mathbf{R}^p$ a $\mathbf{R}$ monótonamente decreciente en el sentido de que si $x \in D$, entonces

$$f_1(x) \geq f_2(x) \geq \cdots \geq f_n(x) \geq f_{n+1}(x) \geq \cdots$$

Si $\lim(f_n(c)) = 0$ para alguna $c \in D$ y $\varepsilon > 0$ demostrar que existen $m \in \mathbf{N}$ y una vecindad U de c tales que si $n \geq m$ y $x \in U \cap D$, entonces $f_n(x) < \varepsilon$.

24.H. Usar el ejercicio anterior para demostrar el siguiente resultado de U. Dini.† Si (f_n) es una sucesión monótona de funciones continuas que converge en cada punto de un conjunto compacto K en $\mathbf{R}^p$ a una función f continua en K, entonces la convergencia es uniforme en K.

24.I. Demostrar, por medio de ejemplos, que el teorema de Dini no es válido si se omite cualquiera de las hipótesis: que K es compacto o que f es continua.

24.J. Demostrar el siguiente teorema de G. Pólya.‡ Si para cada $n \in \mathbf{N}$ la función f_n de $\mathbf{I}$ a $\mathbf{R}$ es monótonamente creciente y si $f(x) = \lim(f_n(x))$ es continua en $\mathbf{I}$ entonces la convergencia es uniforme en $\mathbf{I}$. (Observe que no se está suponiendo que f_n sea continua.)

24.K. Sea (f_n) una sucesión de funciones continuas de $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ y sea $f(x) = \lim(f_n(x))$ para $x \in D$. Demostrar que f es continua en un punto $c \in D$ si y sólo si para cada $\varepsilon > 0$ existen $m \in \mathbf{N}$ y una vecindad U de c tales que si $x \in D \cap U$, entonces $\|f_m(x) - f(x)\| < \varepsilon$.

24.L. Suponga que $f: \mathbf{R} \to \mathbf{R}$ es uniformemente continua en $\mathbf{R}$ y para $n \in \mathbf{N}$, sea $f_n(x) = f(x + 1/n)$ para $x \in \mathbf{R}$. Demostrar que (f_n) converge uniformemente en $\mathbf{R}$ a f.

24.M. Si $f_2(x) = x^2$ para $x \in \mathbf{I}$, ¿qué tan grande debe ser n para que el enésimo polinomio de Bernstein B_n para f_2 satisfaga $|f_2(x) - B_n(x)| \leq 1/1000$ para toda $x \in \mathbf{I}$?

24.N. Si $f_3(x) = x^3$ para $x \in \mathbf{I}$, calcular el enésimo polinomio de Bernstein para f_3. Demostrar directamente que esta sucesión de polinomios converge uniformemente a f_3 en $\mathbf{I}$.

24.O. Diferenciar la ecuación (24.3) una vez con respecto a s y sustituir $s = x, t = 1 - x$ para dar otra variación de la ecuación (24.7).

24.P. Diferenciar la ecuación (24.3) dos veces con respecto a s para dar otra variación de la ecuación (24.8).

24.Q. *(a)* Sea J un intervalo compacto en $\mathbf{R}$ y sean $a \in \mathbf{R}, c \in J$. Dibujar una gráfica de la función $\varphi: J \to \mathbf{R}$ definida por medio de $\varphi(x) = a + \frac{1}{2}(|x - c| + x + c)$.

(b) Demostrar que toda función continua lineal por partes se puede escribir como la suma de un número finito de funciones $\varphi_1, \ldots, \varphi_n$ de la forma dada en la parte *(a)*.

(c) Suponiendo que en cualquier intervalo compacto la función valor absoluto $A(x) = |x|$ es el límite uniforme de una sucesión de polinomios en x, usar la observa-

†ULISSE DINI (1845-1918) estudió y ejerció el magisterio en Pisa. Trabajó en geometría y en análisis, básicamente en series de Fourier.

‡GEORGE POLYA (1887-) nació en Budapest y ejerció el magisterio en Zurich y Stanford. Es muy conocido por sus trabajos en análisis complejo, probabilidad, teoría de números y en la teoría de la inferencia.

ción de la parte *(b)* para dar una demostración del teorema de aproximación de Weierstrass. (El método de esta demostración se debe a Lebesgue.)

24.R. Demostrar que la función $x \mapsto e^x$ en $\boldsymbol{R}$ no es el límite uniforme en $\boldsymbol{R}$ de una sucesión de polinomios. De donde, el teorema de aproximación de Weierstrass puede no ser válido para intervalos infinitos.

24.S. Demostrar que el teorema de aproximación de Weierstrass no es válido para intervalos abiertos acotados.

Sección 25 Límites de funciones

A pesar de que no es posible dar una definición precisa, por lo general se entiende que el "análisis matemático" es la parte de las matemáticas en que se hace uso sistemático de varios conceptos de límite. Si esta afirmación es razonablemente aceptable, podrá parecerle extraño al lector que se haya esperado tanto tiempo para introducir una sección que trate de límites. Existen varias razones por las que se ha hecho esto; la principal es que el análisis elemental trata varios tipos distintos de operaciones de límites. Se ha analizado ya la convergencia de sucesiones y los límites implícitos en el estudio de continuidad. En los siguientes capítulos se verán las operaciones de límites relacionadas con la derivada y la integral. Aun cuando todos estos conceptos de límite son casos especiales de otro más general, el concepto general es de carácter más bien abstracto. Por este motivo, es preferible introducir y analizar los conceptos por separado y no desarrollar primero la idea general de límite y después especificar. Una vez bien entendidos los casos especiales no es difícil comprender el concepto abstracto. Para una excelente explicación de este límite abstracto, véase el artículo de E. J. McShane que aparece en la lista de libros de consulta.

En esta sección nos ocuparemos del límite de una función en un punto y de algunas ligeras variaciones del mismo tema. Por lo general, este tema se estudia antes de continuidad; de hecho, la misma definición de una función continua en ocasiones se expresa en términos de este límite en vez de usar la definición que se dio en la sección 20. Una de las razones por las que se eligió estudiar por separado continuidad de límite es que se habrán de introducir dos definiciones ligeramente distintas del límite de una función en un punto. Debido a que las dos definiciones se usan bastante, se ofrecerán ambas tratando de relacionarlas entre sí.

A menos que se especifique lo contrario, f será una función con dominio D, contenido en $\boldsymbol{R}^p$ y con valores en $\boldsymbol{R}^q$ y se habrá de considerar al límite de f en un punto de acumulación c de D. Por lo tanto, toda vecindad de c contiene una infinidad de puntos de D.

25.1 DEFINICION. *(i)* Se dice que un elemento b de $\boldsymbol{R}^q$ es el **límite supreso de f en c** si para toda vecindad V de b hay una vecindad U de c tal que

si x pertenece a $U \cap D$ y $x \neq c$, entonces $f(x)$ pertenece a V. En este caso se escribe

(25.1) $$b = \lim_c f \qquad \text{ó} \qquad b = \lim_{x \to c} f(x).$$

(ii) Se dice que un elemento b de $\mathbf{R}^q$ es el **límite no supreso de** f **en** c si para toda vecindad V de b hay una vecindad U de c tal que si x pertenece a $U \cap D$, entonces $f(x)$ pertenece a V. En este caso se escribe

(25.2) $$b = \operatorname{Lim}_c f \qquad \text{ó} \qquad b = \operatorname{Lim}_{x \to c} f(x).$$

Observe que la diferencia entre estos dos conceptos se centra en el hecho de que el valor $f(c)$, cuando existe, se tome en cuenta o no. Observe también la diferencia, un tanto sutil, de la notación que se ha introducido en las ecuaciones (25.1) y (25.2). Note que la mayoría de los autores utilizan sólo una de estas notaciones, en cuyo caso se refieren a ella simplemente como "el límite" y por lo general es la notación (25.1) la que usan. Debido a que el límite supreso es el más común se decidió conservar el simbolismo tradicional al referirse a él.

La unicidad de cualquiera de los límites, cuando existe, se plantea con facilidad. Será suficiente la siguiente afirmación.

25.2 LEMA. *(a) Si cualquiera de los límites* $\lim_c f$ *y* $\operatorname{Lim}_c f$ *existe entonces está determinado de manera única.*

(b) Si el límite no supreso existe, entonces el límite supreso existe y

$$\lim_c f = \operatorname{Lim}_c f.$$

(c) Si c no pertenece al dominio D de f, entonces el límite supreso existe si y sólo si el límite no supreso existe.

La parte *(b)* del lema recién planteado prueba que el concepto de límite no supreso es, en cierto modo, más restrictivo que el de límite supreso. La parte *(c)* prueba que los dos límites pueden ser distintos sólo en el caso en que c pertenezca a D. Para dar un ejemplo en que estos dos conceptos difieran, considere la función f de $\boldsymbol{R}$ a $\boldsymbol{R}$ definida por

(25.3) $$\begin{aligned} f(x) &= 0, \qquad x \neq 0, \\ &= 1, \qquad x = 0. \end{aligned}$$

Si $c = 0$, entonces existe el límite supreso de f en $c = 0$ y es igual a 0, mientras que el límite no supreso no existe.

Se establecerán ahora algunas condiciones necesarias y suficientes para la existencia de los límites, dejando las demostraciones para el lector. Se deberá observar que en la parte *(c)* de ambos resultados el límite se refiere al límite de una *sucesión* que se vio en la sección 14.

25.3 TEOREMA. *Las siguientes afirmaciones, relacionadas al límite supreso, son equivalentes.*

(a) El límite supreso $b = \lim_c f$ *existe.*

(b) Si $\varepsilon > 0$, *existe a* $\delta > 0$ *tal que si* $x \in D$ *y* $0 < \|x - c\| < \delta$, *entonces* $\|f(x) - b\| < \varepsilon$.

(c) Si (x_n) *es cualquier sucesión en D tal que* $x_n \neq c$ *y* $c = \lim (x_n)$, *entonces* $b = \lim (f(x_n))$.

25.4 TEOREMA. *Las siguientes afirmaciones, referentes al límite no supreso, son equivalentes.*

(a) El límite no supreso $b = \mathrm{Lim}_c f$ *existe.*

(b) Si $\varepsilon > 0$, *existe* $\delta > 0$ *tal que si* $x \in D$ *y* $\|x - c\| < \delta$, *entonces* $\|f(x) - b\| < \varepsilon$.

(c) Si (x_n) *es cualquier sucesión en D tal que* $c = \lim (x_n)$ *entonces se tiene* $b = \lim (f(x_n))$.

El siguiente resultado da una relación instructiva entre los dos límites y la continuidad de f en c.

25.5. TEOREMA. Si c es un punto de acumulación que pertenece al dominio D de f, entonces las siguientes afirmaciones son equivalentes.

(a) La función f es continua en c.

(b) El límite supreso $\lim_c f$ *existe y es igual a* $f(c)$.

(c) El límite no supreso $\mathrm{Lim}_c f$ *existe*

DEMOSTRACION. Si *(a)* es válido y V es una vecindad de $f(c)$, entonces existe una vecindad U de c tal que si x pertenece a $U \cap D$, entonces $f(x)$ pertenece a V. Es claro que esto implica que Lim f existe en c y es igual a $f(c)$. Análogamente, $f(x)$ pertenece a V para toda $x \neq c$ para la cual $x \in U \cap D$, en cuyo caso lim f existe y es igual a $f(c)$. A la inversa, es fácil ver que las afirmaciones *(b)* y *(c)* implican *(a)*. Q.E.D.

Si f y g son dos funciones que tienen límite supreso (respectivamente, no supreso) en un punto de acumulación c de $D(f+g) = D(f) \cap D(g)$, entonces la suma g tiene límite supreso (respectivamente, no supreso) en c y

$$\lim_c (f+g) = \lim_c f + \lim_c g,$$

$$\left(\text{respectivamente, } \mathrm{Lim}_c (f+g) = \mathrm{Lim}_c f + \mathrm{Lim}_c g\right).$$

Resultados análogos son válidos para otras combinaciones algebraicas de funciones, como se puede ver fácilmente. El siguiente resultado, referente a la composición de dos funciones, es ligeramente más profundo y es un caso en que el límite no supreso es más sencillo que el límite supreso.

25.6 TEOREMA. *Suponga que f tiene dominio $D(f)$ en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$ y que g tiene dominio $D(g)$ en $\boldsymbol{R}^q$ y rango en $\boldsymbol{R}^r$. Sea $g \circ f$ una composición de g y f y sea c un punto de acumulación de $D(g \circ f)$.*

(a) Si ambos límites supresos $b = \lim_c f$ y $a = \lim_b g$ existen y si g es continua en b o bien $f(x) \neq b$ para x en una vecindad de c, entonces el límite supreso de $g \circ f$ existe en c y $a = \lim_c g \circ f$.

(b) Si ambos límites no supresos $b = \text{Lim}_c f$ y $a = \text{Lim}_b g$ existen entonces el límite no supreso de $g \circ f$ existe en c y

$$a = \underset{c}{\text{Lim}}\, g \circ f.$$

DEMOSTRACION. *(a)* Sea W una vecindad de a en $\boldsymbol{R}^r$; dado que $a = \lim g$ en b, hay una vecindad V de b tal que si y pertenece a $V \cap D(g)$ y $y \neq b$ entonces $g(y) \in W$. Como $b = \lim f$ en c, hay una vecindad U de c tal que si x pertenece a $U \cap D(f)$ y $x \neq c$, entonces $f(x) \in V$. Por lo tanto, si x pertenece al conjunto, posiblemente más pequeño, $U \cap D(g \circ f)$, y $x \neq c$, entonces $f(x) \in V \cap D(g)$. Si $f(x) \neq b$ en alguna vecindad U_1 de c, se deduce que para $x \neq c$ en $(U_1 \cap U) \cap D(g \circ f)$, $(g \circ f)(x) \in W$, de tal manera que a es el límite supreso de $g \circ f$ en c. Si g es continua en b, entonces $(g \circ f)(x) \in W$ para $U \cap D(g \circ f)$ y $x \neq c$.

Para demostrar la parte *(b)*, observe que las excepciones hechas en la demostración de *(a)* ya no son necesarias. Por lo que si x pertenece a $U \cap D(g \circ f)$, entonces $f(x) \in V \cap D(g)$ y, por lo tanto, $(g \circ f)(x) \in W$.

Q.E.D.

La conclusión de la parte *(a)* en el teorema anterior puede no ser válida si se omite la condición de que g sea continua en b o que $f(x) \neq b$ en una vecindad de c. Para concretar esta observación, sea f la función de $\boldsymbol{R}$ a $\boldsymbol{R}$ definida en la fórmula (25.3) y sean $g = f$ y $c = 0$. Entonces, $g \circ f$ está dada por

$$\begin{aligned} (g \circ f)(x) &= 1, \qquad x \neq 0, \\ &= 0, \qquad x = 0. \end{aligned}$$

Más aún, se tiene $\lim_{x \to 0} f(x) = 0$, y $\lim_{y \to 0} g(y) = 0$, mientras que es claro que $\lim_{x \to 0} (g \circ f)(x) = 1$. (Observe que los límites no supresos no existen para estas funciones.)

Límites superiores en un punto

En lo que resta de esta sección se habrá de considerar el caso en que $q = 1$. Así pues, f es una función con dominio D en $\boldsymbol{R}^p$ y valores en $\boldsymbol{R}$ y el punto c en $\boldsymbol{R}^p$ es un punto de acumulación de D. Se habrá de definir el límite por arriba o el límite superior de f en c. De nuevo hay dos posibilidades que dependen del hecho de que se consideren vecindades supresas o no supresas, y ambas se habrán de analizar. Es claro que se puede definir el límite inferior de

manera análoga. Una cosa que se debe observar aquí es que, aun cuando la existencia del límite en $\boldsymbol{R}$ (supreso o no) es un asunto relativamente delicado, los límites superiores que se habrán de definir tienen la propiedad de que si f está acotada entonces su existencia está garantizada.

Las ideas que se estudian en esta parte son paralelas al concepto de límite superior de una sucesión en $\boldsymbol{R}^p$ introducido en la sección 19. Sin embargo, no se espera que haya familiarización con lo que allí se vio, excepto en algunos ejercicios.

25.7 DEFINICION. Suponga que f esta acotada en una vecindad del punto c. Si $r>0$, defínanse $\varphi(r)$ y $\Phi(r)$ por medio de

(a) $$\varphi(r) = \sup\{f(x) : 0 < \|x-c\| < r,\ x \in D\},$$

(b) $$\Phi(r) = \sup\{f(x) : \|x-c\| < r,\ x \in D\}$$

y establezca

(c) $$\limsup_{x\to c} f = \inf\{\varphi(r) : r>0\},$$

(d) $$\operatorname{Lim\,sup}_{x\to c} f = \inf\{\Phi(r) : r>0\}.$$

A estas cantidades se las llama el **límite superior supreso** y el **límite superior no supreso** de f en c, respectivamente.

Debido a que estas cantidades están definidas como el ínfimo de la imagen bajo f de vecindades siempre decrecientes de c, probablemente no sea claro que merezcan la terminología "límite superior". En el siguiente lema se da una justificación de esta terminología.

25.8 LEMA. *Si φ, y Φ se definen como antes entonces*

(a) $$\limsup_{x\to c} f = \lim_{r\to 0} \varphi(r),$$

(b) $$\operatorname{Lim\,sup}_{x\to c} f = \lim_{r\to 0} \Phi(r).$$

DEMOSTRACION. Observe que si $0<r<s$, entonces

$$\limsup_{x\to c} f \le \varphi(r) \le \varphi(s).$$

Más aún, por 25.7*(c)*, si $\varepsilon>0$ existe una $r_\varepsilon>0$ tal que

$$\varphi(r_\varepsilon) < \limsup_{x\to c} f + \varepsilon.$$

Por lo tanto, si r satisface $0<r\le r_\varepsilon$, se tiene $|\varphi(r) - \limsup_{x\to c} f| < \varepsilon$, que demuestra *(a)*. La demostración de *(b)* es análoga y se habrá de omitir.

Q.E.D.

25.9 LEMA. *(a) Si $M > \lim \sup_{x \to c} f$, entonces existe una vecindad U de c tal que*

$$f(x) < M \quad \text{para} \quad c \neq x \in D \cap U.$$

(b) Si $M > \text{Lim} \sup_{x \to c} f$, entonces existe una vecindad U de c tal que

$$f(x) < M \quad \text{para} \quad x \in D \cap U.$$

DEMOSTRACION. *(a)* Por 25.7*(c)* se tiene $\inf\{\varphi(r) : r > 0\} < M$. Por lo que existe un número real $r_1 > 0$ tal que $\varphi(r_1) < M$ y se puede tomar $U = \{x \in \mathbf{R}^p : \|x - c\| < r_1\}$. La demostración de *(b)* es análoga. Q.E.D.

25.10 LEMA. *Sean f y g funciones acotadas en una vecindad de c y suponga que c es un punto de acumulación de $D(f+g)$. Entonces*

(a) $$\lim_{x \to c} \sup (f+g) \leq \lim_{x \to c} \sup f + \lim_{x \to c} \sup g$$

(b) $$\text{Lim}_{x \to c} \sup (f+g) \leq \text{Lim}_{x \to c} \sup f + \text{Lim}_{x \to c} \sup g.$$

DEMOSTRACION. Por la relación

$$\sup\{f(x) + g(x) : x \in A\} \leq \sup\{f(x) : x \in A\} + \sup\{g(x) : x \in A\},$$

es claro que usando la notación de la definición 25.7 se tiene

$$\varphi_{f+g}(r) \leq \varphi_f(r) + \varphi_g(r).$$

Ahora usando el lema 25.8 y haciendo $r \to 0$ se obtiene *(a)*. Q.E.D.

Algunos resultados que atañen a otras combinaciones algebraicas se encontrarán en el ejercicio 25.F.

Aunque no se podrá continuar con estos temas, en ciertas áreas del análisis es útil tener la siguiente generalización del concepto de continuidad.

25.11 DEFINICION. Se dice que una función f de D a $\mathbf{R}$ es **semicontinua superior** en un punto c en D siempre que

(25.4) $$f(c) = \text{Lim}_{x \to c} \sup f.$$

Se dice que es **semicontinua superior** en D si es semicontinua superior en todo punto de D.

En vez de definir la semicontinuidad superior en términos de la ecuación (25.4), se podría emplear la condición equivalente pero menos elegante

(25.5) $$f(c) \geq \lim_{x \to c} \sup f.$$

Una de las claves de la importancia y la utilidad de funciones semicontinuas

superiores se puede obtener del siguiente lema que se puede comparar con el teorema de continuidad global 22.1.

25.12 LEMA. *Sea f una función semicontinua superior con dominio D en $\mathbf{R}^p$ y sea k un número real arbitrario. Entonces existen un conjunto abierto G y un conjunto cerrado F tales que*

$$(25.6)\qquad G \cap D = \{x \in D : f(x) < k\}, \qquad F \cap D = \{x \in D : f(x) \geq k\}.$$

DEMOSTRACION. Suponga que c es un punto en D tal que $f(c) < k$. De acuerdo con la definición 25.11 y el lema 25.9*(b)*, hay una vecindad $U(c)$ de c tal que $f(x) < k$ para toda x en $D \cap U(c)$. Sin perder su carácter general, se puede elegir $U(c)$ como una vecindad abierta; estableciendo,

$$G = \bigcup\{U(c) : c \in D\},$$

se tiene un conjunto abierto con la propiedad establecida en (25.6). Si F es el complemento de G, entonces F es cerrado en $\mathbf{R}^p$ y satisface la condición dada. Q.E.D.

Usando el lema que se acaba de demostrar (cf. ejercicio 25.M) es posible probar que si K es un subconjunto compacto de $\mathbf{R}^p$ y f es semicontinua superior en K entonces f está acotada por arriba en K y *existe un punto en K en donde f alcanza su supremo.* De modo que las funciones semicontinuas superiores en conjuntos compactos poseen algunas de las propiedades que se han dado para funciones continuas, aun cuando una función semicontinua superior puede tener muchos puntos de discontinuidad.

El lector se habrá imaginado que es posible extender el concepto de límite superior en un punto al caso en que la función no sea acotada usando ideas de los últimos renglones al final de la sección 18. Análogamente, se puede definir el límite superior como $x \to \pm\infty$. Estas ideas son muy útiles y se dejan como ejercicios.

Ejercicios

25.A. Analizar la existencia de los límites supreso y no supreso de las siguientes funciones en el punto $x = 0$.

(a) $f(x) = |x|$,

(b) $f(x) = 1/x, \quad x \neq 0$,

(c) $f(x) = x \operatorname{sen}(1/x), \quad x \neq 0$,

(d) $f(x) = \operatorname{sen}(1/x), \quad x \neq 0$

(e) $f(x) = \begin{cases} x \operatorname{sen}(1/x), & x \neq 0, \\ 1, & x = 0, \end{cases}$

(f) $f(x) = \begin{cases} 0, & x \leq 0, \\ 1, & x > 0. \end{cases}$

25.B. Demostrar el lema 25.2.

25.C. Si f designa a la función definida en la ecuación (25.3), demostrar que el límite supreso en $x = 0$ es igual a 0 y que el límite no supreso en $x = 0$ no existe. Analizar la existencia de estos dos límites para la composición $f \circ f$.

25.D. Demostrar el lema 25.4.

25.E. Demostrar que las afirmaciones 25.5(*b*) y 25.5(*c*) implican la afirmación 25.5(*a*).

25.F. Demostrar que si f y g tienen límites supresos en un punto de acumulación c del conjunto $D(f) \cap D(g)$, entonces la suma $f+g$ tiene límite supreso en c y

$$\lim_c (f+g) = \lim_c f + \lim_c g.$$

Con la misma hipótesis, el producto interno $f \cdot g$ tiene límite supreso en c y

$$\lim_c (f \cdot g) = \left(\lim_c f\right) \cdot \left(\lim_c g\right).$$

25.G. Sea f la función definida en un subconjunto $D(f)$ en $\boldsymbol{R}$ a $\boldsymbol{R}^q$. Si c es un punto de acumulación del conjunto $V = \{x \in \boldsymbol{R} : x \in D(f),\ x > c\}$, y f_1 es la restricción de f a V, entonces se define al **límite supreso por la derecha de** f **en** c como el $\lim_c f_1$, siempre que este límite exista. Algunas veces el límite se denota por $\lim_{c+} f$ o con $f(c+0)$. Formular y establecer un resultado análogo al teorema 25.3 para el límite supreso por la derecha. (Una definición análoga se puede dar para el **límite no supreso por la derecha** y para ambos límites por la izquierda en c.)

25.H. Sea f una función definida en $D = \{x \in \boldsymbol{R} : x \geq 0\}$ a $\boldsymbol{R}$. Se dice que un número L es el límite de f en $+\infty$ si para cada $\varepsilon > 0$ existe un número real $m(\varepsilon)$ tal que si $x \geq m(\varepsilon)$, entonces $|f(x) - L| < \varepsilon$. En este caso se escribe $L = \lim_{x \to +\infty} f$. Formular y demostrar un resultado análogo al teorema 25.3 para este límite.

25.I. Si f está definida en un conjunto $D(f)$ de $\boldsymbol{R}$ a $\boldsymbol{R}$ y c es un punto de acumulación de $D(f)$, entonces se dice que $f(x) \to +\infty$ conforme $x \to c$, o que

$$\lim_{x \to c} f = +\infty$$

siempre que para cada número positivo M exista una vecindad U de c tal que si $x \in U \cap D(f)$, $x \neq c$, entonces $f(x) > M$. Formular y establecer un resultado análogo al teorema 25.3 para este límite.

25.J. Con base en los ejercicios 25.H y 25.I, dar una definición de lo que significan las expresiones:

$$\lim_{x \to +\infty} f = +\infty, \qquad \lim_{x \to c} f = -\infty.$$

25.K. Probar el lema 25.8 para el límite superior no supreso. Dar la demostración del lema 25.9(*b*).

25.L. Definir lo que significa $\lim\sup_{x \to +\infty} f = L$, y $\lim\inf_{x \to +\infty} f = -\infty$.

25.M. Probar que si f es una función semicontinua superior en un subconjunto compacto K de $\boldsymbol{R}^p$ con valores en $\boldsymbol{R}$ entonces f está acotada por arriba y alcanzan su supremo en K.

25.N. Probar que una función semicontinua superior en un conjunto compacto puede no estar acotada por abajo y puede no alcanzar su ínfimo.

25.O. Probar que si A es un subconjunto abierto de $\boldsymbol{R}^p$ y si f está definida de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ por medio de $f(x) = 1$ para $x \in A$, y $f(x) = 0$ para $x \notin A$, entonces f es una función semicontinua inferior. Si A es un subconjunto cerrado de $\boldsymbol{R}^p$, probar que f es semicontinua superior.

25.P. Dar un ejemplo de una función semicontinua superior que tenga un número infinito de puntos de discontinuidad.

25.Q. ¿Es verdad que una función de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ es continua en un punto si y sólo si es semicontinua superior e inferior en este punto?

25.R. Si (f_n) es una sucesión acotada de funciones continuas de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ y si f^* está definida en $\boldsymbol{R}^p$ por medio de $f^*(x) = \sup \{f_n(x) : n \in \boldsymbol{N}\}$ para $x \in \boldsymbol{R}^p$, ¿entonces es verdad que f^* es semicontinua superior en $\boldsymbol{R}^p$?

25.S. Si (f_n) es una sucesión acotada de funciones continuas de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ y si f_* está definida en $\boldsymbol{R}^p$ por medio de $f_*(x) = \inf \{f_n(x) : n \in \boldsymbol{N}\}$ para $x \in \boldsymbol{R}^p$, ¿entonces es verdad que f_* es semicontinua superior en $\boldsymbol{R}^p$?

25.T. Sea f una función definida en un subconjunto D de $\boldsymbol{R}^p \times \boldsymbol{R}^q$ con valores en $\boldsymbol{R}^r$. Sea (a, b) un punto de acumulación de D. Por analogía con la definición 19.4, definir el límite doble y los dos límites iterados de f en (a, b). Probar que la existencia del límite doble y de los límites iterados implica su igualdad. Probar que el límite doble puede existir sin que exista ninguno de los límites iterados y que los dos límites iterados pueden existir y ser iguales sin que exista el límite doble.

25.U. Sea f como en el ejercicio anterior. Por analogía con las definiciones 17.4 y 19.8, definir lo que significa que

$$g(y) = \lim_{x \to a} f(x, y)$$

uniformemente para y en un conjunto. Formular y probar un resultado análogo al teorema 19.10.

25.V. Sea f como en la definición 25.1 y supóngase que existe el límite supreso en c y que para algún elemento A en $\boldsymbol{R}^q$ y $r > 0$ la desigualdad $\|f(x) - A\| < r$ es válida en alguna vecindad de c. Demostrar que $\|\lim_c f - A\| \le r$. ¿Es válida la misma conclusión para el límite no supreso?

25.W. Analizar la semicontinuidad superior y la inferior de las funciones de los incisos (g) y (h) del ejemplo 20.5.

25.X. Si $f : [0, +\infty) \to \boldsymbol{R}$ es continua en $[0, +\infty)$ y $\lim_{x \to +\infty} f(x) = 0$, probar que f es uniformemente continua en $[0, +\infty)$.

Sección 26 Otros resultados

En esta sección se plantearán algunos teoremas que no se aplicarán posteriormente pero que a menudo son útiles en topología y análisis. Los primeros resultados son extensiones (de largo alcance) del teorema de aproximación de Weierstrass, después hay un teorema que ofrece condiciones en que una función continua tiene una extensión continua y el último resultado es análogo al de Bolzano-Weierstrass en el espacio $C_{pq}(K)$ de funciones continuas en un conjunto compacto K.

El teorema de Stone-Weierstrass

Para facilitar el análisis se introduce la siguiente terminología. Si f y g son funciones con dominio D en $\boldsymbol{R}^p$ y con valores en $\boldsymbol{R}$ entonces las funciones h y k, definidas para x en D por medio de

$$h(x) = \sup \{f(x), g(x)\}, \qquad k(x) = \inf \{f(x), g(x)\},$$

reciben el nombre de **supremo** e **ínfimo,** respectivamente, de las funciones f y g. Si f y g son continuas en D, entonces h y k también son continuas. Esto se deduce del teorema 20.7 y la observación de que si a, b son números reales entonces

$$\sup\{a, b\} = \tfrac{1}{2}\{a + b + |a - b|\},$$
$$\inf\{a, b\} = \tfrac{1}{2}\{a + b - |a - b|\}.$$

En seguida se demuestra un modelo de la generalización de Stone† al teorema de aproximación de Weierstrass. No obstante su reciente descubrimiento, ya se considera "clásico" y debe formar parte de los fundamentos de todo estudiante de matemáticas. El lector deberá consultar el artículo de Stone que aparece en la lista de libros de consulta para obtener extensiones, aplicaciones y un análisis más completo del que aquí se ofrece.

26.1. TEOREMA DE APROXIMACION DE STONE. *Sea K un subconjunto compacto de $\mathbf{R}^p$ y sea $\mathcal{L}$ una colección de funciones continuas de K a $\mathbf{R}$ con las siguientes propiedades:*

(a) Si f y g pertenecen a $\mathcal{L}$, entonces $\sup\{f, g\}$ e $\inf\{f, g\}$ pertenecen a $\mathcal{L}$.

(b) Si $b \in \mathbf{R}$ y $x \neq y \in K$, con x, entonces existe una función f en $\mathcal{L}$ tal que $f(x) = a$, $f(y) = b$.

Entonces, cualquier función continua de K a $\mathbf{R}$ se puede aproximar uniformemente en K por funciones en $\mathcal{L}$.

DEMOSTRACION. Sea F una función continua de K a $\mathbf{R}$. Si x, y pertenecen a K, sea $g_{xy} \in \mathcal{L}$ tal que $g_{xy}(x) = F(x)$ y $g_{xy}(y) = F(y)$. Dado que las funciones F, g_{xy} son continuas y tienen el mismo valor en y; dada $\varepsilon > 0$, hay una vecindad abierta $U(y)$ de y tal que si z pertenece a $K \cap U(y)$, entonces (26.1)

$$g_{xy}(z) > F(z) - \varepsilon. \tag{26.1}$$

Fijando x, para cada $y \in K$, elija una vecindad abierta $U(y)$ con esta propiedad. Por la compacidad de K, se deduce que K está contenido en un número finito de vecindades como: $U(y_1), \ldots, U(y_n)$. Si $h_x = \sup\{g_{xy_1}, \ldots, g_{xy_n}\}$, entonces a partir de la relación (26.1) se deduce que (26.2)

$$h_x(z) > F(z) - \varepsilon \qquad \text{for } z \in K. \tag{26.2}$$

Dado que $g_{xy_j}(x) = F(x)$, se puede ver que $h_x(x) = F(x)$ y, por lo tanto, hay una vecindad abierta $V(x)$ de x tal que si z pertenece a $K \cap V(x)$, entonces

†MARSHALL H. STONE (1903-) estudió en Harvard y ha sido profesor ahí y en las universidades de Chicago y Massachusetts. Hijo de un presidente de la Suprema Corte, ha hecho aportaciones básicas al análisis moderno, especialmente a las teorías del espacio de Hilbert y de las álgebras booleanas.

(26.3) $$h_x(z) < F(z) + \varepsilon.$$

Usar una vez más la compacidad de K para obtener un número finito de vecindades $V(x_1), \ldots, V(x_m)$ y sea $h = \inf\{h_{x_1}, \ldots, h_{x_m}\}$ Entonces h pertenece a $\mathscr{L}$ y se deduce de (26.2) que

$$h(z) > F(z) - \varepsilon \quad \text{para } z \in K.$$

y de (26.3) que

$$h(z) < F(z) + \varepsilon \quad \text{para } z \in K.$$

Combinando estos resultados se tiene $|h(z) - F(z)| < \varepsilon$, $z \in K$, que da la aproximacion deseada. Q.E.D.

El lector se podrá haber dado cuenta de que en este último resultado no se utilizó el teorema de aproximación de Weierstrass. En el siguiente resultado se sustituye la condición *(a)* por tres condiciones algebraicas en el conjunto de funciones. En esta ocasión se utiliza el clásico teorema de Weierstrass 24.8 para el caso especial de la función valor absoluto φ definida para t en $\boldsymbol{R}$ como $\varphi(t) = |t|$, para concluir que φ se puede aproximar por polinomios en todo conjunto compacto de números reales.

26.2 TEOREMA DE STONE-WEIERSTRASS. Sea K un subconjunto compacto de $\boldsymbol{R}^p$ y sea $\mathscr{A}$ una colección de funciones continuas de K a $\boldsymbol{R}$ con las siguientes propiedades:

(a) La función constante $e(x) \to 1$, $x \in K$, pertenece a $\mathscr{A}$.

(b) Si f, g pertenecen a $\mathscr{A}$, entonces $\alpha f + \beta g$ pertenecen a $\mathscr{A}$ para todas α, β en $\boldsymbol{R}$.

(c) Si f, g pertenecen a $\mathscr{A}$, entonces fg pertenece a $\mathscr{A}$.

(d) Si $x \neq y$ son dos puntos de K, existe una función f en $\mathscr{A}$ tal que $f(x) \neq f(y)$.

Entonces, cualquier función continua de K a $\boldsymbol{R}$ se puede aproximar uniformemente en K por funciones en $\mathscr{A}$.

DEMOSTRACION. Sean $a, b \in \boldsymbol{R}$ y $x \neq y$ tales que pertenezcan a K. De acuerdo con *(d)*, hay una función f en $\mathscr{A}$ tal que $f(x) \neq f(y)$. Dado que $e(x) = 1 = e(y)$, se deduce que hay números reales α, β tales que

$$\alpha f(x) + \beta e(x) = a, \qquad \alpha f(y) + \beta e(y) = b.$$

Por lo tanto, por *(b)* existe una función $g \in \mathscr{A}$ tal que $g(x) = a$ y $g(y) = b$.

Ahora, sea $\mathscr{L}$ el conjunto de todas las funciones continuas en K que se pueden aproximar uniformemente por funciones en $\mathscr{A}$. Es obvio que $\mathscr{A} \subseteq \mathscr{L}$, de tal manera que $\mathscr{L}$ tiene la propiedad *(b)* del teorema de aproximación de Stone 26.1. Se demostrará ahora que si $h \in \mathscr{L}$, entonces $|h| \in \mathscr{L}$. Dado que

$$\sup\{f, g\} = \tfrac{1}{2}(f + g + |f - g|),$$
$$\inf\{f, g\} = \tfrac{1}{2}(f + g - |f - g|),$$

esto implicará que $\mathcal{L}$ tiene la propiedad 26.1*(a)* y por lo tanto que toda función continua de K a $\boldsymbol{R}$ pertenece a $\mathcal{L}$.

Como h es continua y K es compacto, se deduce que existe una $M > 0$ tal que $\|h\|_K \le M$. Dado que $h \in \mathcal{L}$, hay una sucesión (h_n) de funciones en $\mathcal{A}$ que convergen uniformemente a h en K y se puede suponer que $\|h_n\|_K \le M + 1$ para toda $n \in \boldsymbol{N}$. (¿por qué?) Si $\varepsilon > 0$ está dada, se aplica el teorema de aproximación de Weierstrass 24.8 a la función valor absoluto en el intervalo $[-(M+1), M+1]$ para obtener un polinomio p_ε tal que

$$\big|\,|t| - p_\varepsilon(t)\big| \le \tfrac{1}{2}\varepsilon \quad \text{para } |t| \le M + 1.$$

Se deduce por lo tanto que

$$\big|\,|h_n(x)| - p_\varepsilon(h_n(x))\big| \le \tfrac{1}{2}\varepsilon \quad \text{para } x \in K.$$

Pero $p_\varepsilon \circ f_h$ pertenece a $\mathcal{A}$ debido a las hipótesis (a), (b) y (c). Como

$$\big|\,|h(x)| - |h_n(x)|\,\big| \le \|h - h_n\|_K$$

se deduce que si n es suficientemente grande entonces se tiene

$$\big|\,|h(x)| - p_\varepsilon \circ h(x)\big| \le \varepsilon \quad \text{para } x \in K.$$

Dado que $\varepsilon > 0$ es arbitrario, se deduce que $|h| \in \mathcal{L}$ y ahora el resultado se deduce del teorema anterior. Q.E.D.

Como caso especial del teorema de Stone-Weierstrass se obtendrá una forma más fuerte del teorema 24.8. Este resultado es más fuerte que el anterior por dos razones: (i) permite que el dominio sea un subconjunto compacto arbitrario de $\boldsymbol{R}^p$ y no únicamente una celda compacta en $\boldsymbol{R}$ y (ii) permite que el rango esté en cualquier espacio $\boldsymbol{R}^q$, y no sólo en $\boldsymbol{R}$. Para entender el planteamiento, recuerde que una función f con dominio D en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$ se puede considerar como q funciones de D a $\boldsymbol{R}$ con la representación de coordenadas:

(26.4) $$f(x) = (f_1(x), \ldots, f_q(x)) \quad \text{para } x \in D.$$

Si cada función coordenada f_j es un polinomio en las p coordenadas $(x_1, \ldots, x_p)$, entonces se dice que f es una función polinomial.

26.3 TEOREMA DE APROXIMACION POLINOMIAL. *Sea f una función continua cuyo dominio K es un subconjunto compacto de $\boldsymbol{R}^p$ y cuyo rango pertenece a $\boldsymbol{R}^q$ y sea $\varepsilon > 0$. Entonces existe una función polinomial p de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ tal que $\|f(x) - p(x)\| < \varepsilon$ para $x \in K$.*

DEMOSTRACION. Representar a f por medio de sus q funciones coordenadas, como en (26.4). Dado que f es continua en K, cada una de las funciones coordenadas f_j es continua de K a $\boldsymbol{R}$. Es evidente que las funciones polinomiales definidas de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ satisfacen las propiedades del teorema de Stone-Weierstrass. Por lo tanto, la función coordenada f_j se puede aproximar uniformemente en K en $\varepsilon/\sqrt{q}$ por una función polinomial p_j. Estando p definida por

$$p(x) = (p_1(x), \ldots, p_q(x)),$$

se obtiene una función polinomial de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ que da la aproximación deseada en K a la función dada f. Q.E.D.

Extensión de funciones continuas

Algunas veces es deseable extender el dominio de una función continua a un conjunto más grande sin cambiar los valores del dominio original. Esto siempre se puede hacer de una manera trivial estableciendo que la función sea 0 fuera del dominio original, pero, en general, este método de extensión no da una función continua. Después de reflexionar un poco, el lector podrá ver que no siempre es posible obtener una extensión continua. Por ejemplo, si $D = \{x \in \boldsymbol{R} : x \neq 0\}$ y si f está definida para $x \in D$ como $f(x) = 1/x$, entonces no es posible extender a f de tal manera que se obtenga una función continua en todo $\boldsymbol{R}$. Sin embargo, es importante saber que una extensión siempre es posible cuando el dominio es un conjunto cerrado. Además, no es necesario incrementar la cota de la función (si es que es acotada).

Antes de demostrar este teorema de extensión, observe que si A y B son dos subconjuntos cerrados ajenos de $\boldsymbol{R}^p$, entonces existe una función continua φ definida en $\boldsymbol{R}^p$ con valores de $\boldsymbol{R}$ tal que

$$\varphi(x) = 0, \quad x \in A; \qquad \varphi(x) = 1, \quad x \in B; \qquad 0 \le \varphi(x) \le 1, \quad x \in \boldsymbol{R}^p.$$

De hecho, si $d(x, A) = \inf \{\|x - y\| : y \in A\}$ y $d(x, B) = \inf \{\|x - y\| : y \in B\}$, entonces se puede definir φ para $x \in \boldsymbol{R}^p$ por medio de la ecuación

$$\varphi(x) = \frac{d(x, A)}{d(x, A) + d(x, B)}.$$

26.4 TEOREMA DE EXTENSION DE TIETZE†. *Sea f una función continua acotada definida en un subconjunto cerrado D de $\boldsymbol{R}^p$ y con valores en $\boldsymbol{R}$. Entonces existe una función continua g de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ tal que $g(x) = f(x)$ para x en D y tal que* $\sup \{\|g(x)\| : x \in \boldsymbol{R}^p\} = \sup \{\|f(x)\| : x \in D\}$.

†HEINRICH TIETZE (1880-1964) fue profesor en Munich e investigador en topología, geometría y álgebra. Este teorema de extensión data de 1914.

DEMOSTRACION. Sea $M = \sup\{|f(x)| : x \in D\}$ y considere $A_1 = \{x \in D : f(x) \leq -M/3\}$ y $B_1 = \{x \in D : f(x) \geq M/3\}$. Por la continuidad de f y el hecho de que D sea cerrado, del teorema 22.1(c) se deduce que A_1 y B_1 son subconjuntos cerrados de $\boldsymbol{R}^p$. De acuerdo con la observación que precede al planteamiento del teorema, hay una función continua φ_1 de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ tal que

$$\varphi_1(x) = -\tfrac{1}{3}M, \quad x \in A_1; \qquad \varphi_1(x) = \tfrac{1}{3}M, \quad x \in B_1;$$
$$-\tfrac{1}{3}M \leq \varphi_1(x) \leq \tfrac{1}{3}M, \quad x \in \boldsymbol{R}^p.$$

Fije $f_2 = f - \varphi_1$ y observe que f_2 es continua en D y que $\sup\{|f_2(x)| : x \in D\} \leq \tfrac{2}{3}M$.

Siguiendo, se definen $A_2 = \{x \in D : f_2(x) \leq -\tfrac{1}{3}\tfrac{2}{3}M\}$ $B_2 = \{x \in D : f_2(x) \geq \tfrac{1}{3}\tfrac{2}{3}M\}$ y se obtiene una función continua φ_2 de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ tal que

$$\varphi_2(x) = -\tfrac{1}{3}\tfrac{2}{3}M, \quad x \in A_2; \qquad \varphi_2(x) = -\tfrac{1}{3}\tfrac{2}{3}M, \quad x \in B_2;$$
$$-\tfrac{1}{3}\tfrac{2}{3}M \leq \varphi_2(x) \leq \tfrac{1}{3}\tfrac{2}{3}M, \quad x \in \boldsymbol{R}^p.$$

Una vez que se haga esto, fije $f_3 = f_2 - \varphi_2$ y observe que $f_3 = f - \varphi_1 - \varphi_2$ es continua en D y que $\sup\{|f_3(x)| : x \in D\} \leq (\tfrac{2}{3})^2 M$.

Siguiendo de esta manera se obtiene una sucesión (φ_n) de funciones definidas de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ tales que, para cada n,

$$(26.5) \qquad |f(x) - [\varphi_1(x) + \varphi_2(x) + \cdots + \varphi_n(x)]| \leq (\tfrac{2}{3})^n M,$$

para toda x en D y tal que

$$(26.6) \qquad |\varphi_n(x)| \leq (\tfrac{1}{3})(\tfrac{2}{3})^{n-1} M \qquad \text{para } x \in \boldsymbol{R}^p.$$

Defínase g_n de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ por medio de $g_n = \varphi_1 + \varphi_2 + \cdots + \varphi_n$, por lo que se deduce que g_n es continua. De la desigualdad (26.6) se deduce que si $m \geq n$ y $x \in \boldsymbol{R}^p$, entonces

$$|g_m(x) - g_n(x)| = |\varphi_{n+1}(x) + \cdots + \varphi_m(x)| \leq (\tfrac{1}{3})(\tfrac{2}{3})^n M[1 + \tfrac{2}{3} + (\tfrac{2}{3})^2 + \cdots] \leq (\tfrac{2}{3})^n M,$$

lo cual prueba que la sucesión (g_n) converge uniformemente en $\boldsymbol{R}^p$ a una función que se denotará por g. Dado o.ue cada g_n es continua en $\boldsymbol{R}^p$, entonces el teorema 24.1 implica que g es continua en todo punto de $\boldsymbol{R}^p$. Además, por la desigualdad (26.5) se puede ver que $|f(x) - g_n(x)| \leq (\tfrac{2}{3})^n M$ para $x \in D$. Por lo tanto, se concluye que $f(x) = g(x)$ para x in D. Por último, la desigualdad (26.6) implica que para cualquier x en $\boldsymbol{R}^p$ se tiene

$$|g_n(x)| \leq \tfrac{1}{3}M[1 + \tfrac{2}{3} + \cdots + (\tfrac{2}{3})^{n-1}] \leq M,$$

que prueba la última afirmación del teorema. Q.E.D.

26.5 COROLARIO. *Sea f una función continua acotada definida en un subconjunto cerrado D de $\mathbf{R}^p$ y con valores en $\mathbf{R}^q$. Entonces existe una función continua g de $\mathbf{R}^p$ a $\mathbf{R}^q$ con $g(x) = f(x)$ para x en D y tal que*

$$\sup\{\|g(x)\| : x \in \mathbf{R}^p\} \leq \sqrt{q} \sup\{\|f(x)\| : x \in D\}.$$

DEMOSTRACION. Este resultado se acaba de probar para $q = 1$. Para el caso general, observe que f define a q funciones coordenadas continuas de valor real en D, digamos

$$f(x) = (f_1(x), f_2(x), \ldots, f_q(x)).$$

Dado que cada una de las f_j, $1 \leq j \leq q$ tiene una extensión continua g_j de $\mathbf{R}^p$ a $\mathbf{R}$, se define g de $\mathbf{R}^p$ a $\mathbf{R}^q$ por medio de $g(x) = (g_1(x), g_2(x), \ldots, g_q(x))$. Se puede ver que la función g tiene las propiedades requeridas. Q.E.D.

Equicontinuidad

Con frecuencia se ha usado el teorema de Bolzano-Weierstrass 10.6 para conjuntos (que asegura que todo subconjunto acotado infinito de $\mathbf{R}^p$ tiene un punto de acumulación) y el teorema 16.4 equivalente para sucesiones (que asegura que toda sucesión acotada en $\mathbf{R}^p$ tiene una subsucesión convergente). En seguida se presenta un teorema enteramente análogo al teorema Bolzano-Weierstrass pero para *conjuntos de funciones continuas* y no para conjuntos de puntos. Para que sea más sencillo y breve, se ofrecerá sólo una forma secuencial del teorema.

En lo sucesivo, K será un subconjunto compacto fijo de $\mathbf{R}^p$, y se estará tratando con funciones continuas en K con rango en $\mathbf{R}^q$. Por el teorema 22.5, cada una de estas funciones es acotada y por lo tanto $C_{pq}(K) = BC_{pq}(K)$. Se dice que un conjunto $\mathcal{F}$ en $C_{pq}(K)$ es **acotado** (o **uniformemente acotado**) en K si existe una constante M tal que $\|f\|_K \leq M$, para toda f en $\mathcal{F}$. Es claro que cualquier conjunto finito $\mathcal{F}$ de dichas funciones es acotado, ya que si $\mathcal{F} = \{f_1, f_2, \ldots, f_n\}$, entonces se puede fijar

$$M = \sup\{\|f_1\|_K, \|f_2\|_K, \ldots, \|f_n\|_K\}.$$

En general, un conjunto infinito de funciones continuas de K a $\mathbf{R}^q$ no será acotado. Sin embargo, una sucesión uniformemente convergente de funciones continuas es acotada. (Cf. ejercicio 26.M).

Si f es una función continua en un conjunto compacto K de $\mathbf{R}^p$, entonces el teorema 23.3 implica que es uniformemente continua. Por lo tanto, si $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$, tal que x, y pertenecen a K y $\|x - y\| < \delta(\varepsilon)$, entonces $\|f(x) - f(y)\| < \varepsilon$. Desde luego, el valor de δ puede depender de la función f, así como de ε por lo que a menudo se escribe $\delta(\varepsilon, f)$. (Cuando se está

tratando con más de una función, es conveniente indicar explícitamente esta dependencia.) Observe que si $\mathcal{F} = \{f_1, \ldots, f_n\}$ es un conjunto finito en $C_{pq}(K)$, entonces, fijando

$$\delta(\varepsilon, \mathcal{F}) = \inf \{\delta(\varepsilon, f_1), \ldots, \delta(\varepsilon, f_n)\},$$

se obtiene una δ que "sirve" para todas las funciones de este conjunto finito.

26.6 DEFINICION. Se dice que un conjunto de funciones de K a $\mathbf{R}^q$ es **uniformemente equicontinuo** en K si para cada número real $\varepsilon > 0$ hay un número $\delta(\varepsilon) > 0$ tal que si x, y pertenecen a K y $\|x - y\| < \delta(\varepsilon)$ y f es una función de $\mathcal{F}$, entonces $\|f(x) - f(y)\| < \varepsilon$.

Se ha visto que un conjunto finito de funciones continuas en K es equicontinuo. También es cierto que una sucesión de funciones continuas que converge uniformemente en K también es equicontinua. (Cf. ejercicio 26.N.)

Para que una sucesión en $C_{pq}(K)$ sea uniformemente convergente en K es necesario que la sucesión sea acotada y uniformemente equicontinua en K. Se habrá de probar ahora que estas dos propiedades son necesarias y suficientes para que un conjunto $\mathcal{F}$ en $C_{pq}(K)$ tenga la propiedad de que toda sucesión de funciones de $\mathcal{F}$ tiene una subsucesión que converge uniformemente en K. Esto se puede considerar como una generalización del teorema de Bolzano-Weierstrass para conjuntos de funciones continuas y desempeña un papel importante en la teoría de ecuaciones diferenciales e integrales.

26.7 TEOREMA DE ARZELA-ASCOLI†. *Sea K un subconjunto compacto de $\mathbf{R}^p$ y sea $\mathcal{F}$ una colección de funciones continuas en K y con valores en $\mathbf{R}^q$. Las siguientes propiedades son equivalentes:*

(a) La familia $\mathcal{F}$ es acotada y uniformemente equicontinua en K.

(b) Toda sucesión de $\mathcal{F}$ tiene una subsucesión que es uniformemente convergente en K.

DEMOSTRACION. Primero se probará que si la condición *(a)* es falsa, entonces también lo es la condición *(b)*. Si $\mathcal{F}$ no es acotada, entonces existe una sucesión (f_n) en $\mathcal{F}$ tal que $\|f_n\|_K \geq n$ para $n \in \mathbf{N}$. Pero entonces ninguna subsucesión de (f_n) puede ser uniformemente convergente. Además, si el conjunto $\mathcal{F}$ no es uniformemente equicontinuo, entonces para alguna $\varepsilon_0 > 0$ existe (¿por qué?) una sucesión (f_n) en $\mathcal{F}$ y sucesiones (x_n) y (y_n) en K con $\|x_n - y_n\| < 1/n$ pero tales que $\|f_n(x_n) - f_n(y_n)\| > \varepsilon_0$. Pero entonces ninguna subsucesión de (f_n) puede ser uniformemente convergen en K.

†CESARE ARZELA (1847-1912) fue profesor en Bolonia. Dio las condiciones necesarias y suficientes para que el límite de una sucesión de funciones continuas en un intervalo cerrado sea continuo y estudió otros temas en relación con esto.

GIULIO ASCOLI (1843-1912), profesor en Milán, formuló la definición de equicontinuidad en un planteamiento geométrico. También hizo aportaciones en series de Fourier.

Se demostrará ahora que si el conjunto $\mathcal{F}$ satisface (a) entonces dada cualquier sucesión (f_n) en $\mathcal{F}$ hay una subsucesión que converge uniformemente en K. Para hacer esto, observe que a partir del ejercicio 10.H se deduce que existe un conjunto numerable C en K tal que si $y \in K$ y $\varepsilon > 0$, entonces existe un elemento x en C tal que $\|x - y\| < \varepsilon$. Si $C = \{x_1, x_2, \ldots\}$, entonces la sucesión $(f_n(x_1))$ es acotada en $\boldsymbol{R}^q$ del teorema de Bolzano-Weierstrass 16.4 se deduce que hay una subsucesión

$$(f_1{}^1(x_1), f_2{}^1(x_1), \ldots, f_n{}^1(x_1), \ldots)$$

de $(f_n(x_1))$ que es convergente. Observe ahora que la sucesión $(f_k{}^1(x_2) : k \in \boldsymbol{N})$ es acotada en $\boldsymbol{R}^q$; por tanto, tiene una subsucesión

$$(f_1{}^2(x_2), f_2{}^2(x_2), \ldots, f_n{}^2(x_2), \ldots)$$

que es convergente. Nuevamente, la sucesión $(f_n{}^2(x_3) : n \in \boldsymbol{N})$ es acotada en $\boldsymbol{R}^q$, por lo que alguna subsucesión

$$(f_1{}^3(x_3), f_2{}^3(x_3), \ldots, f_n{}^3(x_3), \ldots)$$

es convergente. Se prosigue de la misma manera y después se fija $g_n = f_n{}^n$ de tal manera que g_n sea la enésima función en la enésima subsucesión. Por la construcción, es claro que la sucesión (g_n) converge en cada punto de C.

Se probará ahora que la sucesión (g_n) converge en cada punto de K y que la convergencia es uniforme. Para hacer esto, sea $\varepsilon > 0$ y sea $\delta(\varepsilon)$ como en la definición 26.6. Sea $C_1 = \{y_1, \ldots, y_k\}$ un subconjunto finito de C tal que todo punto en K dista menos de $\delta(\varepsilon)$ algún punto en C_1 dado que las sucesiones

$$(g_n(y_1)), (g_n(y_2)), \ldots, (g_n(y_k))$$

convergen, exite un número natural M tal que si $m, n \geq M$, entonces

$$\|g_m(y_i) - g_n(y_i)\| < \varepsilon \quad \text{para } i = 1, 2, \ldots, k.$$

Dada $x \in K$, existe una $y_j \in C_1$ tal que $\|x - y_j\| < \delta(\varepsilon)$ de donde, por la equicontinuidad uniforme, se tiene $\|g_n(x) - g_n(y_j)\| < \varepsilon$ para toda $n \in \boldsymbol{N}$; específicamente, esta desigualdad es válida para $n \geq M$. Por lo tanto, se tiene

$$\begin{aligned}\|g_n(x) - g_m(x)\| &\leq \|g_n(x) - g_n(y_j)\| + \|g_n(y_j) - g_m(y_j)\| \\ &\quad + \|g_m(y_j) - g_m(x)\| < \varepsilon + \varepsilon + \varepsilon = 3\varepsilon,\end{aligned}$$

siempre que $m, n \geq M$. Esto prueba que

$$\|g_n - g_m\|_K \leq 3\varepsilon \quad \text{para } m, n \geq M,$$

por lo que la convergencia uniforme de la sucesión (g_n) en K se sigue del criterio de Cauchy para convergencia uniforme dado en 17.11. Q.E.D.

En la demostración de este resultado se construyó una sucesión de subsucesiones de funciones y después se releccionó la sucesión "diagonal" (g_n), en donde $g_n = f_n^n$. A dicha construcción con frecuencia se le llama "proceso diagonal" o "método diagonal de Cantor" y con frecuencia es útil. El lector recordará que un tipo similar de argumento se usó en la sección 3 para demostrar que los números reales no forman un conjunto contable.

Ejercicios

26.A. Demostrar que la condición *(a)* del teorema 26.1 es equivalente a la condición:

(a') Si f pertenece a $\mathcal{L}$, entonces $|f|$ pertenece a $\mathcal{L}$.

26.B. Demostrar que toda función continua de valor real, en el intervalo $[0, \pi]$ es el límite uniforme de una sucesión de "polinomios en cos x" (es decir, de funciones (P_n) en donde $P_n(x) = p_n(\cos x)$ para algún polinomio p_n).

26.C. Demostrar que toda función continua de valor real en $[0, \pi]$ es el límite uniforme de una sucesión de funciones de la forma

$$x \mapsto a_0 + a_1 \cos x + a_2 \cos 2x + \cdots + a_n \cos nx.$$

26.D. Explicar por qué el resultado del ejercicio 26.B no es válido cuando $\cos kx$ se reemplaza por sen kx, $k \in \boldsymbol{N}$.

26.E. Usar el ejercicio 26.C para probar que toda función continua de valor real f en $[0, \pi]$ con $f(0) = f(\pi)$ es el límite uniforme de una sucesión de funciones de la forma

$$x \mapsto b_0 + b_1 \operatorname{sen} x + b_2 \operatorname{sen} 2x + \cdots + b_n \operatorname{sen} nx.$$

26.F. Usar los ejercicios 26.C y 26.E para probar que toda función continua de valor real f en $[-\pi, \pi]$ con $f(-\pi) = f(\pi)$ es el límite uniforme de una sucesión de funciones de la forma

$$x \mapsto a_0 + a_1 \cos x + b_1 \operatorname{sen} x + \cdots + a_n \cos nx + b_n \operatorname{sen} nx.$$

[*Sugerencia:* dividir a f en la suma $f = f_e + f_o$ de una función par $f_e(x) = \frac{1}{2}(f(x) + f(-x))$ y de una función impar $f_o(x) = \frac{1}{2}(f(x).]$

26.G. Dar una demostración para el ejercicio anterior basado en el teorema 26.3 aplicado al círculo unitario $T = \{(x, y) \in \boldsymbol{R}^2 : x^2 + y^2 = 1\}$ y las observaciones de que hay una correspondencia uno a uno entre funciones continuas de T a $\boldsymbol{R}$ y funciones continuas de $[-\pi, \pi]$ a $\boldsymbol{R}$ que satisfacen $f(-\pi) = f(\pi)$.

26.H. Sea $J \subseteq \boldsymbol{R}$ un intervalo compacto y sea $\mathcal{A}$ una colección de funciones continuas en $J \to \boldsymbol{R}$ que satisfacen las propiedades del teorema de Stone-Weierstrass 26.2. Probar que cualquier función continua en $J \times J$ (in $\boldsymbol{R}^2$) a $\boldsymbol{R}$ se puede aproximar uniformemente por funciones de la forma

$$f_1(x)g_1(y) + \cdots + f_n(x)g_n(y).$$

en donde f_i, g_i pertenecen a $\mathcal{A}$.

26.I. Probar que el teorema de Tietze 26.4 puede no ser válido si el dominio no es cerrado.

26.J. Usar el teorema de Tietze 26.4 para probar que si $D \subseteq \boldsymbol{R}^p$ es cerrado y si f es una función continua no acotada en $D \to \boldsymbol{R}$ entonces existe una extensión continua de f a todo $\boldsymbol{R}^p$. [*Sugerencia:* considere la composición $\phi \circ f$, en donde $\phi(x) =$ Arc tan x $\phi(x) = x/(1+x)$.]

26.K. Sea $\mathcal{F}$ una colección de funciones de $D \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^q$. Considere la siguiente propiedad en el punto $c \in D$: si $\varepsilon > 0$, hay una $\delta(c, \varepsilon) > 0$ tal que si $x \in D$ y $\|x - c\| < \delta(c, \varepsilon)$ entonces $\|f(x) - f(c)\| < \varepsilon$ para toda $f \in \mathcal{F}$. Probar que $\mathcal{F}$ tiene esta propiedad en $c \in D$ si y sólo si para cada sucesión (x_n) en D con $c = \lim (x_n)$, entonces $f(c) = \lim (f(x_n))$ uniformemente para $f \in \mathcal{F}$. (Algunas veces se dice que $\mathcal{F}$ es **equicontinua en** $c \in D$ cuando se cumpla esta propiedad.)

26.L. Sea $\mathcal{F}$ igual que en ejercicio 26.K. si D es compacto y se satisface la propiedad del ejercicio 26.K. Para toda $c \in D$, probar que $\mathcal{F}$ es uniformemente equicontinua en el sentido de la definición 26.6.

26.M. Si $K \subseteq \boldsymbol{R}^p$ es compacto y (f_n) es una sucesión de funciones continuas de K a $\boldsymbol{R}^q$ uniformemente convergente en K, probar que la familia $\{f_n\}$ es acotada en K (en el sentido de que existe $M > 0$ tal que $\|f_n(x)\| \leq M$ para toda $x \in K$, $n \in \boldsymbol{N}$ o $\|f_n\|_K \leq M$ para $n \in \boldsymbol{N}$).

26.N. Si $K \subseteq \boldsymbol{R}^p$ es compacto y (f_n) es una sucesión de funciones continuas de K a $\boldsymbol{R}^q$ uniformemente convergente en K, probar que la familia $\{f_n\}$ es uniformemente equicontinua en K en el sentido de la definición 26.6.

26.O. Sea $\mathcal{F}$ una colección de funciones, acotada y uniformemente equicontinua, de $D \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}$ y defínase a f^* en $D \to \boldsymbol{R}$ por medio de

$$f^*(x) = \sup \{f(x) : f \in \mathcal{F}\}.$$

Probar que f^* es continua de D a $\boldsymbol{R}$.

26.P. Demostrar que la conclusión del ejercicio anterior puede no ser válida si no se supone que $\mathcal{F}$ sea uniformemente equicontinua.

26.Q. Considere las siguientes sucesiones de funciones, las cuales prueban que el teorema de Arzela-Ascoli 26.7 puede no ser válido si se eliminan las distintas hipótesis.

(a) $f_n(x) = x + n$ para $x \in [0, 1]$;
(b) $f_n(x) = x^n$ para $x \in [0, 1]$;
(c) $f_n(x) = \dfrac{1}{1 + (x - n)^2}$ para $x \in [0, +\infty)$.

26.R. Sea (f_n) una sucesión de funciones continuas de $\boldsymbol{R}$ a $\boldsymbol{R}^q$ que converge en cada punto del conjunto Q de números racionales. Si el conjunto $\{f_n\}$ es uniformemente equicontinuo en $\boldsymbol{R}$, probar que la sucesión converge en todo punto de $\boldsymbol{R}$ y que la convergencia es uniforme en todo conjunto compacto de $\boldsymbol{R}$ pero no necesariamente uniforme en $\boldsymbol{R}$.

V
FUNCIONES DE UNA VARIABLE

Se iniciará el estudio de diferenciación e integración de funciones. Para ello será conveniente considerar primero el caso de funciones de *una* variable; en los capítulos VII y VIII se volverán a ver funciones de más de una variable. Al comparar con estos capítulos se podrá ver que el caso de funciones con más de una variable es bastante parecido, en términos generales, a lo que se verá aquí; sin embargo, surgen ciertas complicaciones. Además, dado que en la teoría general se usan resultados del caso de una variable, es conveniente haber estudiado previamente este caso.

En las secciones 27 y 28 se introduce la derivada de una función definida en un intervalo real y se demuestran el importante teorema del valor medio y algunos de sus corolarios. En la sección 29 se introduce la definición de la integral de Riemann (y Riemann-Stieltjes) de funciones acotadas en un intervalo $[a, b]$. Las propiedades básicas de la integral se demuestran en esta sección y en las secciones 30 y 31, en ellas se analizan las integrales "impropia" e infinita. Aun cuando los resultados de estas secciones se usan muy poco en las siguientes partes del libro, son importantes en muchas aplicaciones.

Sección 27. Teorema del valor medio

Como se supone que el lector ya está familiarizado con la relación que hay entre la derivada de una función de $\boldsymbol{R}$ a $\boldsymbol{R}$ y la pendiente de su gráfica y con el concepto de razón de cambio (instantánea), se concentra la atención por completo en los aspectos matemáticos de la derivada y no se verán las aplicaciones en física, economía, etc. En esta sección y en la siguiente se habrá de considerar una función con dominio D y rango contenido en $\boldsymbol{R}$. Aun cuando el principal interés está en la derivada en un punto interior, se definirá la derivada de una manera más general de manera que se pueda considerar, por ejemplo, al punto extremo de un intervalo. Sin embargo, sí se requiere que el punto en el que se está definiendo la derivada sea punto de acumulación de D y pertenezca a D.

27.1 DEFINICION. Si c es punto de acumulacion de D y pertenece a D, se dice que un número real L es la **derivada de f en c** si para todo número $\varepsilon > 0$ hay un número $\delta(\varepsilon) > 0$ tal que si x pertenece a D y $0 < |x-c| < \delta(\varepsilon)$ entonces

$$\left| \frac{f(x)-f(c)}{x-c} - L \right| < \varepsilon. \tag{27.1}$$

En este caso se escribe $f'(c)$ como L.

Alternativamente, se podría definir a $f'(c)$ como el límite

$$\lim_{x \to c} \frac{f(x)-f(c)}{x-c} \qquad (x \in D, x \neq c).$$

Se debe observar que si c es un punto interior de D entonces en (27.1) se considera a los puntos x tanto a la izquierda como a la derecha del punto c. Por otro lado, si D es un intervalo y c es el punto extremo izquierdo de D, entonces en (27.1) sólo se puede tomar x a la derecha de c.

Siempre que exista la derivada de f en c su valor se denota por medio de $f'(c)$. De esta manera se obtiene una función f' cuyo dominio es un subconjunto del dominio de f. En seguida se demuestra que la continuidad de f en c es una condición necesaria para la existencia de la derivada en c.

27.2 LEMA. *Si f tiene una derivada en c, entonces f es continua en ese punto.*

DEMOSTRACION. Sea $\varepsilon = 1$ y tome $\delta = \delta(1)$ tal que

$$\left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < 1,$$

para toda $x \in D$ que satisfaga $0 < |x-c| < \delta$. Por la desigualdad del triángulo se deduce que para estos valores de x se tiene

$$|f(x)-f(c)| \leq |x-c| \{|f'(c)|+1\}.$$

El lado izquierdo de esta expresión se puede hacer menor que ε si se toma x en D con $|x-c| < \inf\{\delta, \varepsilon/(|f'(c)|+1)\}$.

Q.E.D.

Es fácil ver que la continuidad en c no es una condición suficiente para que exista la derivada en c. Por ejemplo, si $D = \boldsymbol{R}$ y $f(x) = |x|$, entonces f es continua en todo punto de $\boldsymbol{R}$ pero tiene derivada en el punto c si y sólo si $c \neq 0$. Usando combinaciones algebraicas sencillas es fácil construir funciones continuas que no tengan alguna derivada en un número finito, e incluso contable, de puntos. En 1872 Weierstrass revolucionó el mundo de los matemáticos dando un ejemplo de una *función que es continua en todo punto pero cuya derivada no existe en ninguna parte.* (De hecho la función definida por la serie

$$f(x) = \sum_{n=0}^{\infty} \frac{1}{2^n} \cos(3^n x),$$

se puede probar que tiene esta propiedad. No se verán los detalles pero se recomiendan al lector los libros Titchmarsh y Boas para más información y como referencia.)

27.3 LEMA. *(a) Si f tiene una derivada en c y $f'(c)>0$, existe un número $\delta>0$ tal que si $x \in D$ y $c<x<c+\delta$, entonces $f(c)<f(x)$.*

(b) Si $f'(c)<0$, existe un número $\delta>0$ tal que $x \in D$ y $c-\delta<x<c$, entonces $f(c)<f(x)$.

DEMOSTRACION *(a)* Sea ε_0 tal que $0<\varepsilon_0<f'(c)$ y tome $\delta=\delta(\varepsilon_0)$ dependiendo de ε_0 como en la definición 27.1. Si $x \in D$ y $c<x<c+\delta$, entonces se tiene

$$-\varepsilon_0<\frac{f(x)-f(c)}{x-c}-f'(c).$$

Dado que $x-c>0$, esta relación implica que

$$0<(f'(c)-\varepsilon_0)(x-c)<f(x)-f(c),$$

lo cual prueba la afirmación *(a)*. La demostración de *(b)* es análoga
Q.E.D.

Recuerde que se dice que una función f tiene un **máximo relativo** en un punto c en D si existe una $\delta>0$ tal que $f(x) \leq f(c)$ cuando $x \in D$ satisface $|x-c|<\delta$. Una definición análoga es aplicable al término **mínimo relativo.** El siguiente resultado proporciona la justificación teórica del conocido proceso de encontrar puntos en los que f tenga máximos y mínimos relativos, analizando los ceros de la derivada. Observe que este procedimiento sólo es aplicable a puntos interiores del intervalo. De hecho, si $f(x)=x$ en $D=[0, 1]$, entonces el punto extremo $x=0$ da al único mínimo relativo y el punto extremo $x=1$ da al único máximo relativo de f; sin embargo, ninguno es raíz de la derivada. Para hacerlo más sencillo, se probará este resultado sólo para máximos relativos, dejando al lector la formulación de los resultados correspondientes para mínimos relativos.

27.4 TEOREMA DEL MAXIMO INTERIOR. *Sea c un punto interior de D en donde f tiene un máximo relativo. Si existe la derivada de f en c, entonces debe ser igual a cero.*

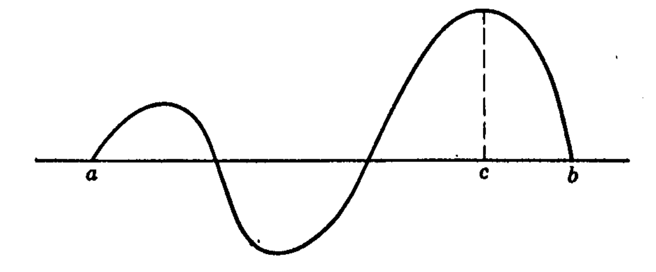


Figura 27.1

DEMOSTRACION. Si $f'(c) > 0$, entonces por el lema 27.3*(a)* hay una $\delta > 0$ tal que si $c < x < c + \delta$ y $x \in D$, entonces $f(c) < f(x)$. Esto contradice el supuesto de que f tenga un máximo relativo en c. Si $f'(c) < 0$, se usa el lema 27.3*(b)*.

Q.E.D.

27.5 TEOREMA DE ROLLE.† *Suponga que f es continua en un intervalo cerrado $J = [a, b]$, que la derivada f' existe en el intervalo abierto y que $f(a) = f(b) = 0$. Entonces existe un punto c en (a,b) tal que $f'(c) = 0$.*

DEMOSTRACION Si f es constante en J, se puede tom[illegible] $c = (a + b)/2$. Por lo que se va a suponer que f no es constante; substituyendo f por $-f$, si es que es necesario, se puede suponer que f toma algunos valores positivos. Por el teorema del valor máximo 22.7, la función f adquiere el valor sup $\{f(x) : x \in J\}$ en algún punto c de J. Dado que $f(a) = f(b) = 0$, el punto c satisface $a < c < b$. Por hipótesis, $f'(c)$ existe y, dado que f tiene un punto máximo relativo en c, el teorema del máximo interior implica que $f'(c) = 0$.

Como consecuencia del teorema de Rolle, se obtiene el teorema del valor medio.

27.6 TEOREMA DEL VALOR MEDIO. *Suponga que f es continua en un intervalo cerrado $J = [a, b]$ y que tiene una derivada en el intervalo abierto (a, b). Entonces existe un punto c en (a, b) tal que*

$$f(b) - f(a) = f'(c)(b - a).$$

DEMOSTRACION. Considere la función φ definida en J por

$$\varphi(x) = f(x) - f(a) - \frac{f(b) - f(a)}{b - a}(x - a).$$

(Es fácil ver que φ es la diferencia entre f y la función cuya gráfica consta del segmento de línea que pasa por los puntos $(a, f(a))$ y $(b, f(b))$; véase la figura 27.2) De la hipótesis se deduce que φ es continua en $J = [a, b]$ y es fácil comprobar que φ tiene una derivada en (a, b). Más aún, se tiene $\varphi(a) = \varphi(b) = 0$. Aplicando el teorema de Rolle,
existe un punto c dentro de J tal que

$$0 = \varphi'(c) = f'(c) - \frac{f(b) - f(a)}{b - a}$$

de donde se deduce el resultado

†Este teorema generalmente se le atribuye a MICHEL ROLLE (1652-1719), miembro de la Academia Francesa, quien hizo aportaciones a la geometría analítica y a los primeros trabajos que dieron origen al cálculo.

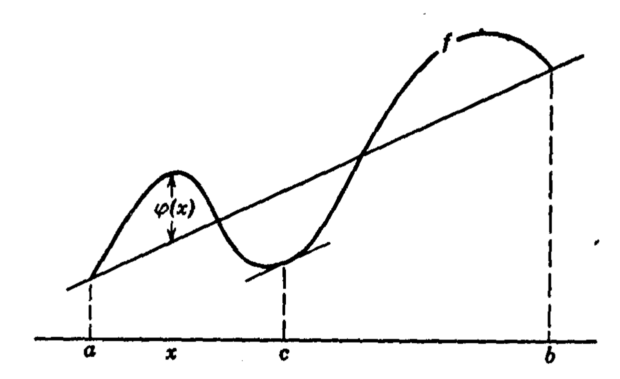


Figura 27.2 El teorema del valor medio.

27.7 COROLARIO. *Si f iene una derivada en $J=[a, b]$, entonces existe un punto c en (a, b) tal que*

$$f(b)-f(a)=f'(c)(b-a).$$

En algunas ocasiones es conveniente tener una versión más general del teorema del valor medio para dos funciones.

27.8 TEOREMA DE CAUCHY DEL VALOR MEDIO. *Sean f, g continuas en $J=[a, b]$ y con derivadas dentro de (a, b). Entonces existe un punto c en (a, b) tal que*

$$f'(c)[g(b)-g(a)]=g'(c)[f(b)-f(a)].$$

DEMOSTRACION Cuando $g(b)=g(a)$ el resultado es inmediato si se toma a c de tal manera que $g'(c)=0$. Si $g(b)\neq g(a)$, considere la función φ definida en J por medio de

$$\varphi(x)=f(x)-f(a)-\frac{f(b)-f(a)}{g(b)-g(a)}[g(x)-g(a)].$$

Aplicando el teorema de Rolle a φ, se obtiene el resultado que se esperaba.

Q.E.D.

Aun cuando la derivada de una función *no necesariamente* es continua, existe un teorema elemental pero notable debido a Darboux† que asegura que la derivada f' alcanza todo valor entre $f'(a)$ y $f'(b)$ en el intervalo $[a, b]$. (Véase el ejercicio 27.H.)

Es fácil recordar el enunciado del teorema del valor medio dibujando diagramas apropiados. Esto no se debe descartar aun tiende a sugerir que su importancia es de naturaleza geométrica, lo que es un tanto erróneo De hecho, el teorema del valor medio es un lobo con piel de oveja y es *el* teorema

† GASTON DARBOUX (1842-1917) fue alumno de Hermite y profesor en el Colegio de Francia. A pesar de ser conocido primordialmente como geómetra, también hizo importantes aportaciones al análisis.

fundamental del cálculo diferencial. Se cierra esta sección con algunas consecuencias elementales de dicho resultado. Mas resultados se darán en la siguiente sección, e incluso aparecerán aún más adelante.

27.9 TEOREMA. *Supóngase que f es continua en* $J=[a, b]$ *y que su derivada existe en* (a, b).

(i) Si $f'(x)=0$ *para* $a<x<b$, *entonces f es constante en J.*

(ii) Si $f'(x)=g'(x)$ *para* $a<x<b$, *entonces f y g difieren en J por una constante.*

(iii) Si $f'(x)\geq 0$ *para* $a<x<b$ *y si* $x_1\leq x_2$ *pertenecen a J, entonces* $f(x_1)\leq f(x_2)$.

(iv) Si $f'(x)>0$ *para* $a<x<b$ *y si* $x_1<x_2$ *pertenecen a J, entonces* $f(x_1)<f(x_2)$.

(v) Si $f'(x)\geq 0$ *para* $a<x<a+\delta$, *entonces a es un punto mínimo relativo de f.*

(vi) Si $f'(x)\geq 0$ *para* $b-\delta<x<b$, *entonces b es un punto máximo relativo de f.*

(vii) Si $|f'(x)|\leq M$ *para* $a<x<b$, *entonces f satisface la condición de Lipschitz:*

$$|f(x_1)-f(x_2)|\leq M\,|x_1-x_2| \quad \text{paia } x_1, x_2 \text{ in } J.$$

La demostración se le deja al lector.

Ejercicios

27.A Usando la definición, calcular la derivada (cuando exista) de las funciones dadas por las expresiones:

(a) $f(x)=x^2$ para $x\in \mathbf{R}$,
(b) $g(x)=x^n$ para $x\in \mathbf{R}$,
(c) $h(x)=\sqrt{x}$ para $x\geq 0$,
(d) $F(x)=1/x$ para $x\neq 0$,
(e) $G(x)=|x|$ para $x\in \mathbf{R}$,
(f) $H(x)=1/x^2$ para $x\neq 0$.

27.B. Si f y g son funciones de valor real definidas en en un intervalo J y si son diferenciales en un punto c, demostrar que su producto h, definido como $h(x)=f(x)g(x)$, para $x\in J$, es diferenciable en c y

$$h'(c)=f'(c)g(c)+f(c)g'(c).$$

27.C. Probar que la función definida para $x\neq 0$ por medio de

$$f(x)=\operatorname{sen}(1/x)$$

es diferenciable en cada número real distinto de cero. Probar que su derivada no está acotada en una vecindad de $x=0$. (Se puede hacer uso de identidades trigonométricas, la continuidad de las funciones seno y coseno y la relación elemental de límite $(\operatorname{sen} u)/u \to 1$ cuando $u\to 0$.)

27.D Demostrar que la función definida por

$$g(x) = x^2 \operatorname{sen}(1/x), \qquad x \neq 0,$$
$$= 0, \qquad x = 0,$$

es diferenciable para todos los números reales, pero g' no es continua en $x = 0$.

27.E. La función $h: \mathbf{R} \to \mathbf{R}$ definida por medio de $h(x) = x^2$ para $x \in \mathbf{Q}$ y $h(x) = 0$ para $x \notin \mathbf{Q}$ es continua exactamente en un punto. ¿Es diferenciable en ese punto?

27 F. Sea $c \in D$ un punto de acumulación de D y sea $f: D \to \mathbf{R}$. probar que $f'(c)$ existe si y sólo si para toda sucesión (x_n) en D con $x_n \neq c$ para $n \in \mathbf{N}$ tal que $(x_n) = c$, el límite de la sucesión

$$\left(\frac{f(x_n) - f(c)}{x_n - c}\right)$$

existe. En este caso los límites de todas estas sucesiones son iguales a $f'(c)$.

27.G. Si $f: D \to \mathbf{R}$ es diferenciable en $c \in D$ y si $c + 1/n \in D$ para toda $n \in \mathbf{N}$, demostrar que

$$f'(c) = \lim\left(n\{f(c + 1/n) - f(c)\}\right).$$

Sin embargo, demostrar que la existencia del límite de esta sucesión no implica la existencia de la derivada.

27.H. (Darboux) Si f es diferenciable en $[a, b]$, si $f'(a) = A$, $f'(b) = B$ y C está entre A y B, entonces existe un punto c en (a, b) para el cual $f'(c) = C$. (*Sugerencia:* considere la cota inferior de la función $g(x) = f(x) - C(x - a)$.)

27.I. Si $g(x) = 0$ para $x < 0$ y $g(x) = 1$ para $x \geq 1$, demostrar que no existe una función $f: \mathbf{R} \to \mathbf{R}$ tal que $f'(x) = g(x)$ para toda $x \in \mathbf{R}$.

27 J. Dar un ejemplo de una función continua con un único punto máximo relativo pero tal que la derivada no exista en ese punto.

27.K. Dar un ejemplo de una función uniformemente continua que sea diferenciable en $(0,1)$ pero tal que su derivada no sea acotada en $(0,1)$.

27.L. Sea $f:[a, b] \to \mathbf{R}$ diferenciable en $c \in [a, b]$. Demostrar que si para toda $\varepsilon > 0$, existe una $\delta(\varepsilon) > 0$ tal que si $0 < |x - y| < \delta(\varepsilon)$ y $a \leq x \leq c \leq y \leq b$, entonces

$$\left|\frac{f(x) - f(y)}{x - y} - f'(c)\right| < \varepsilon.$$

27.M. Sea $f:[a, b] \to \mathbf{R}$ diferenciable en $[a, b]$. Probar que f' es continua en $[a, b]$ si y sólo si para toda $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si $0 < |x - y| < \delta(\varepsilon)$, $x, y \in [a, b]$, entonces

$$\left|\frac{f(x) - f(y)}{x - y} - f'(x)\right| < \varepsilon.$$

27.N. Sea $f:[a, b] \to \mathbf{R}$ continua en $[a, b]$ y diferenciable en (a, b). Si $\lim_a f'(x) = A$ demostrar que $f'(a)$ existe y es igual a A.

27.O. Si $f: \mathbf{R} \to \mathbf{R}$ y $f'(a)$ existen, demostrar que

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a - h)}{2h}.$$

Sin embargo, dar un ejemplo para probar que la existencia de este límite no implica la existencia de la derivada.

27.P. Se dice que una función $f:\mathbf{R}\to\mathbf{R}$ es **par** si $f(-x)=f(x)$ para toda $x\in\mathbf{R}$, y que es **impar** si $f(-x)=-f(x)$ para toda $x\in\mathbf{R}$. Si f es diferenciable en $\mathbf{R}$ y par (impar, respectivamente), probar que f' es impar (par, respectivamente).

27.Q. Sean $f:(a,b)\to\mathbf{R}$ y $c\in(a,b)$. Se escribe $f(c+)=\lim_{x\to c}f(x)$(el límite por la derecha de f en c). Si el límite por la derecha

$$A_r=\lim_{\substack{x\to c\\ x>c}}\frac{f(x)-f(c+)}{x-c}$$

existe en $\mathbf{R}$, se dice que f tiene una **derivada por la derecha** en c y se denota a A_r por medio de $f'_+(c)$. Análogamente para las derivadas por la izquierda.

Probar que si f es continua en c entonces $f'(c)$ existe si y sólo si $f'_+(c)$ y $f'_-(c)$ existen y son iguales. Demostrar que se puede tener $g'_-(c)=g'_+(c)$ sin que exista $g'(c)$.

27.R. Sean I y J intervalos en $\mathbf{R}$ y sean $f:I\to\mathbf{R}$ y $g:J\to\mathbf{R}$ tales que g sea diferenciable en un punto $b\in J$ y f sea diferenciable en un punto interior $a=g(b)$ de I. Probar que la composición $h=f\circ g$ definida para $\{x\in J: g(x)\in I\}$ es diferenciable en b y que $h'(b)=f'(a)g'(b)$. (*Sugerencia:* defina a H en $D(h)$ por medio de

$$
\begin{aligned}
H(x)&=\frac{f(g(x))-f(g(b))}{g(x)-g(b)} &&\text{si}\quad g(x)\neq g(c),\\
&=f'(a) &&\text{si}\quad g(x)=g(c).
\end{aligned}
$$

Demostrar que $\lim_b H(x)=f'(a)$. Después usar el hecho de que $(g(x)-g(b))H(x)=f(g(x))-f(g(b))$ para toda x en $D(h)$.)

27.S. Sea $f:[0,+\infty)\to\mathbf{R}$ diferenciable en $(0,+\infty)$.

(*a*) Si $f'(x)\to b\in\mathbf{R}$ cuando $x\to+\infty$, probar que para cualquier $h>0$ se tiene

$$\lim_{x\to+\infty}\frac{f(x+h)-f(x)}{h}=b.$$

(*b*) Si $f(x)\to a\in\mathbf{R}$ y $f'(x)\to b\in\mathbf{R}$ cuando $x\to+\infty$, entonces $b=0$.

(*c*) Si $f'(x)\to b\in\mathbf{R}$ cuando $x\to+\infty$, entonces $f(x)/x\to b$ cuando $x\to+\infty$.

27.T. Sea $f:[a,b]\to\mathbf{R}$ diferenciable con $0<m\le f'(x)\le M$ para $x\in[a,b]$ y sea $f(a)<0<f(b)$. Dada $x_1\in[a,b]$, defina la sucesión (x_n) por medio de

$$x_{n+1}=x_n-\frac{1}{M}f(x_n),\qquad n\in\mathbf{N}.$$

Demostrar que esta sucesión está bien definida y que converge a la raíz única $\bar{x}$ de la ecuación $f(x)=0$ en (a,b) y que

$$|x_{n+1}-\bar{x}|\le\frac{|f(x_1)|}{m}\left(1-\frac{m}{M}\right)^n$$

para $n\in\mathbf{N}$. (*Sugerencia:* sea $\varphi:[a,b]\to\mathbf{R}$ tal que $\varphi(x)=x-f(x)/M$. Probar que φ es creciente y una contracción (véase 23.4) con constante $1-m/M$.)

27.U. Sea $f:\mathbf{R}\to\mathbf{R}$ una función que tiene una derivada continua y tal que $f(a)=b$ y $f'(a)\neq0$. Sea $\delta>0$ tal que si $|x-a|\le\delta$ entonces $|f'(x)-f'(a)|\le\frac{1}{2}|f'(a)|$, y

sea $\eta = \frac{1}{2}\delta\,|f'(a)|$. Demostrar que si $|\bar{y} - b| \le \eta$, entonces la sucesión (x_n) definida por $x_1 = a$ y

$$x_{n+1} = x_n - \frac{f(x_n) - \bar{y}}{f'(a)}, \qquad n \in \mathbf{N}$$

converge al punto único $\bar{x}$ en $[a - \delta, a + \delta]$ tal que $f(\bar{x}) = \bar{y}$. (*Sugerencia:* probar que la función definida por $\varphi(x) = x - (f(x) - \bar{y})/f'(a)$ es una contracción con constante $\frac{1}{2}$ en el intervalo $[a - \delta, a + \delta]$.)

Sección 28 Otras aplicaciones del teorema del valor medio

Es difícil que se pueda sobreenfatizar la importancia del teorema del valor medio ya que desempeña un papel primordial en muchas consideraciones teóricas. De igual manera, es sumamente útil en muchos aspectos prácticos. En 27.9 se dieron algunas consecuencias inmediatas del teorema del valor medio que a menudo son útiles. En seguida se proponen otras áreas en las que se puede aplicar; al hacer esto, se hará uso, con más libertad que antes, de la experiencia y los conocimientos del lector en lo que respecta a ciertas funciones bien conocidas.

28.1 APLICACION. El teorema de Rolle se puede usar para localizar raíces de una función. Puesto que si una función g se puede identificar como la derivada de una función f, entonces entre cualesquiera dos raíces de f hay cuando menos una raíz de g. Por ejemplo, sea $g(x) = \cos x$; se sabe que g es la derivada de $f(x) = \operatorname{sen} x$. Por lo tanto, entre cualesquiera dos raíces de sen x existe cuando menos una raíz de cos x. Por otro lado $g'(x) = -\operatorname{sen} x = -f(x)$, y otra aplicación del teorema de Rolle nos dice que entre cualesquiera dos raíces de cos x hay cuando menos una raíz de sen x. Por lo tanto, se concluye que *las raíces* de sen x y cos x se entrelazan entre sí. Es probale que esta conclusión no sea novedad para el lector; sin embargo, el mismo tipo de argumento es aplicable a las *funciones de Bessel*† J_n de orden $n = 0, 1, 2, \ldots$ usando las relaciones

$$[x^n J_n(x)]' = x^n J_{n-1}(x), [x^{-n} J_n(x)]' = -x^{-n} J_{n+1}(x) \qquad \text{para } x > 0.$$

Los detalles de este argumento los deberá proporcionar el lector.

28.2 APLICACION. El teorema del valor medio se puede aplicar para cálculos aproximados y para obtener errores de estima. Por ejemplo, supóngase que se desea calcular $\sqrt{105}$. Se usa el teorema del valor medio con $f(x) = \sqrt{x}$, $a = 100$, $b = 105$ para obtener

†FRIEDRICH WILHELM BESSEL (1784-1846) fue astrónomo y matemático. Amigo cercano de Gauss, se le conoce principalmente por la ecuación diferencial que lleva su nombre.

i230 (1478x2443x2 pcx)

$$\sqrt{105}-\sqrt{100}=\frac{5}{2\sqrt{c}},$$

para algún número c con $100<c<105$. Dado que $10<\sqrt{c}<\sqrt{105}<\sqrt{121}=11$, se puede asegurar que

$$\frac{5}{2(11)}<\sqrt{105}-10<\frac{5}{2(10)},$$

por lo que se deduce que $10.22<\sqrt{105}<10.25$. Este cálculo puede no ser tan preciso como se quisiera. Es claro que el calculo $\sqrt{c}<\sqrt{105}<\sqrt{121}$ fue exagerado y se puede mejorar haciendo uso de la conclusión de que $\sqrt{105}<10.25$. De modo que $\sqrt{c}<10.25$ y fácilmente se determina que

$$0.243<\frac{5}{2(10.25)}<\sqrt{105}-10.$$

El cálculo obtenido es $10.243<\sqrt{105}<10.250$ y cálculos más precisos se pueden obtener de esta manera.

28.3 APLICACION. El teorema del valor medio y sus corolarios se pueden usar para probar desigualdades y para extender desigualdades conocidas para valores enteros o racionales a valores reales.

Por ejemplo, recuerde que la desigualdad de Bernoulli 5.C asegura que si $1+x>0$ y $n\in \boldsymbol{N}$ entonces $(1+x)^n \geq 1+nx$. Se habrá de probar que esta desigualdad es válida para cualquier exponente real $r\geq 1$. Para hacer esto, sea $f(x)=(1+x)^r$, de tal manera que $f'(x)=r(1+x)^{r-1}$. Si $-1<x<0$, entonces $f'(x)<r$, mientras que si $x>0$, entonces $f'(x)>r$. Si se aplica el teorema del valor medio a estos dos casos se obtiene el resultado

$$(1+x)^r \geq 1+rx,$$

cuando $1+x>0$ Más aún, si $r\geq 1$. entonces existe la igualdad si y sólo si $x=0$.

Como resultado análogo, sea α un número real que satisface $0<\alpha<1$ y sea $g(x)=\alpha x-x^\alpha$ para $x\geq 0$. Entonces $g'(x)=\alpha(1-x^{\alpha-1})$, de tal manera que $g'(x)<0$ para $0<x<1$ y $g'(x)>0$ para $x>1$. Como consecuencia, si $x\geq 0$, entonces $g(x)\geq g(1)$ y $g(x)=g(1)$ si y sólo si $x=1$. Por lo tanto, si $x\geq 0$ y $0<\alpha<1$, entonces se tiene

$$x^\alpha \leq \alpha x+(1-\alpha).$$

Si $a\geq 0$ y $b>0$ y si $x=a/b$ se multiplica por b se obtiene la desigualdad

$$a^\alpha b^{1-\alpha} \leq \alpha a+(1-\alpha)b.$$

en donde la igualdad es válida si y sólo si $a=b$. Esta desigualdad con frecuencia es el punto de partida al demostrar la importante desigualdad de Hölder (cf. proyecto $8.\beta$).

28.4 APLICACION. Las conocidas reglas de L'Hospital[†] para el cálculo de "formas indeterminadas" se pueden probar por medio del teorema del valor medio de Cauchy. Por ejemplo, suponga que f, g son continuas en $[a, b]$ y tienen derivadas en (a, b), que $f(a) = g(a) = 0$, pero que g, g' no desaparecen para $x \neq a$. Entonces existe punto c con $a < c < b$ tal que

$$\frac{f(b)}{g(b)} = \frac{f'(c)}{g'(c)}.$$

Se deduce que si existe $\lim_{x \to a} f'(x)/g'(x)$ entonces

$$\lim_{x \to a} \frac{f(x)}{g(x)} = \lim_{x \to a} \frac{f'(x)}{g'(x)}.$$

El caso en el que las funciones llegan a ser infinito en $x = a$, o en el que el punto en que se toma el límite es infinito, o en el que se tiene un "indeterminante." o alguna otra forma a menudo se puede tratar tomando logaritmos, exponenciales o alguna otra forma análoga.

Por ejemplo, si $a = 0$ y se desea calcular el límite de $h(x) = x \log x$ cuando $x \to 0$, no se puede aplicar el argumento anterior. Se escribe $h(x)$ de la forma $f(x)/g(x)$ en donde $f(x) = \log x$ y $g(x) = 1/x$, $x > 0$. Se puede ver que

$$\frac{f'(x)}{g'(x)} = \frac{\frac{1}{x}}{\frac{-1}{x^2}} = -x \to 0, \quad \text{cuando} \quad x \to 0.$$

Sea $\varepsilon > 0$ y elija un número fijo $0 < x_1 < 1$ tal que si $0 < x < x_1$, entonces $|f'(x)/g'(x)| < \varepsilon$. Aplicando el teorema del valor medio de Cauchy se tiene

$$\left| \frac{f(x) - f(x_1)}{g(x) - g(x_1)} \right| = \left| \frac{f'(x_2)}{g'(x_2)} \right| < \varepsilon,$$

en donde x_2 satisface $0 < x < x_2 < x_1$. Dado que $f(x) \neq 0$ y $g(x) \neq 0$ para $0 < x < x_1$, se puede escribir la cantidad que aparece del lado izquierdo de manera más conveniente como

$$\frac{f(x)}{g(x)} \left\{ \frac{1 - \frac{f(x_1)}{f(x)}}{1 - \frac{g(x_1)}{g(x)}} \right\}.$$

† GUILLAUME FRANCOIS L'HOSPITAL (1661-1704) fue alumno de Johann Bernoulli (1667-1748). El Marqués de L'Hospital publicó sus trabajos de cátedra sobre cálculo diferencial en 1696, presentado así, el primer libro de texto sobre cálculo, al mundo.

Manteniendo a x_1 fijo, se deja $x \to 0$. Dado que la cantidad entre corchetes converge a 1, excede a $\frac{1}{2}$ para x lo suficientemente pequeña. De lo anterior se deduce que

$$|h(x)| = \left| \frac{f(x)}{g(x)} \right| < 2\varepsilon,$$

para x suficientemente cerca de 0. Por lo que el límite en $x = 0$ de h es 0.

Intercambio de límite y derivada

Sea (f_n) una sucesión de funciones definida en un intervalo J de $\boldsymbol{R}$ y con valores en $\boldsymbol{R}$. Es fácil dar un ejemplo de una sucesión de funciones que tengan derivadas en todo punto de J y que converja en J a una función f que no tenga derivada en algunos puntos de J. (Hacerlo.) Más aún, el ejemplo de Weierstrass mencionado antes se puede usar para dar un ejemplo de una sucesión de funciones que derivadas en todo punto de $\boldsymbol{R}$ y que converjan uniformemente en $\boldsymbol{R}$ a una función continua que no tenga derivada en ningún punto. De modo que, en general, no es lícito diferenciar el límite de una sucesión convergente de funciones que tengan derivadas aun cuando la convergencia sea uniforme.

Se habrá de probar ahora que si la *sucesión de derivadas* es uniformemente convergente entonces todo está bien. Si se agrega la hipótesis de que las derivadas sean continuas entonces es posible dar una demostración corta con base en la integral de Riemann. Sin embargo, si no se supone que las derivadas sean continuas, es necesario un argumento un poco delicado.

28.5 TEOREMA. *Sea (f_n) una sucesión de funciones definidas en un intervalo finito J de $\boldsymbol{R}$ y con valores en $\boldsymbol{R}$. Suponga que hay un punto x_0 en J en el que la sucesión $(f_n(x_0))$ converge, que las derivadas f'_n existen en J y que la sucesión (f'_n) converge uniformemente en J a una función g. Entonces, la sucesión (f_n) converge uniformemente en J a una función f que tiene una derivada en todo punto de J y $f' = g$.*

DEMOSTRACION. Suponga que los puntos terminales de J son $a < b$ y sea x cualquier punto de J. Si m, n son números naturales, se aplica el teorema del valor medio a la diferencia $f_m - f_n$ en el intervalo con puntos terminales x_0, x para concluir que existe un punto y (dependiendo de m, n) tal que

$$f_m(x) - f_n(x) = f_m(x_0) - f_n(x_0) + (x - x_0)\{f'_m(y) - f'_n(y)\}.$$

De donde se deduce que

$$\|f_m - f_n\|_J \leq |f_m(x_0) - f_n(x_0)| + (b - a)\,\|f'_m - f'_n\|_J$$

de modo que la sucesión (f_n) converge uniformemente en J a una función que se denotará por medio de f. Dado que las f_n son continuas y la convergencia de (f_n) a f es uniforme, entonces f es continua en J.

Para probar la existencia de la derivada de f en un punto c en J se aplica el teorema del valor medio a la diferencia $f_m - f_n$ en un intervalo con puntos terminales c,x para deducir que existe un punto z (dependiendo de m,n) tal que

$$\{f_m(x) - f_n(x)\} - \{f_m(c) - f_n(c)\} = (x - c)\{f'_m(z) - f'_n(z)\}.$$

Se deduce que cuando $c \neq x$, entonces

$$\left| \frac{f_m(x) - f_m(c)}{x - c} - \frac{f_n(x) - f_n(c)}{x - c} \right| \leq \|f'_m - f'_n\|_J.$$

Debido a la convergencia uniforme de la sucesión (f'_n), el lado derecho está dominado por ε cuando $m, n \geq M(\varepsilon)$. Tomando el límite con respecto a m, del lema 15.8 se deduce que

$$\left| \frac{f(x) - f(c)}{x - c} - \frac{f_n(x) - f_n(c)}{x - c} \right| \leq \varepsilon,$$

cuando $n \geq M(\varepsilon)$. Dado que $g(c) = \lim (f'_n(c))$, existe una $N(\varepsilon)$ tal que si $n \geq N(\varepsilon)$ entonces $|f'_n(c) - g(c)| < \varepsilon$. Ahora, sea $K = \sup\{M(\varepsilon), N(\varepsilon)\}$. Por la existencia de $f'_K(c)$, si $0 < |x - c| < \delta_K(\varepsilon)$, entonces

$$\left| \frac{f_K(x) - f_K(c)}{x - c} - f'_K(c) \right| < \varepsilon.$$

Por lo tanto, se deduce que si $0 < |x - c| < \delta_K(\varepsilon)$, entonces

$$\left| \frac{f(x) - f(c)}{x - c} - g(c) \right| < 3\varepsilon.$$

Esto prueba que $f'(c)$ existe y es igual a $g(c)$ Q.E.D.

Teorema de Taylor

Si la derivada $f'(x)$ de f existe en todo punto x del conjunto D, se puede considerar la existencia de la derivada de la función f' en un punto $c \in D$. Cuando f' tenga una derivada en c, al número que resulta se le llama la **segunda derivada** de f en c y normalmente se denota como $f''(c)$, o como $f^{(2)}(c)$. De manera análoga se define la tercera $f'''(c) = f^{(3)}(c), \ldots$, y la n-ésima derivada $f^{(n)}(c), \ldots$, siempre que estas derivadas existan.

En seguida se obtiene el importante teorema que se atribuye a Brook Taylor† que desempeña un papel importante en muchas investigaciones y que se puede considerar como una extensión del teorema del valor medio.

28.6 TEOREMA DE TAYLOR. *Suponga que n es un número natural, que f y sus derivadas $f', f'', \ldots, f^{(n-1)}$ están definidas y son continuas en*

† **BROOK TAYLOR (1685-1731) fue uno de los primeros matemáticos ingleses. En 1715 dio la expansión de series infinitas, pero, siendo fiel al espíritu de su época, no analizó la convergencia. Lo restante lo hizo Lagrange.**

$J=[a,b]$, y que existe en (a,b). Si α, β pertenecen a J, entonces existe un número γ entre $f^{(n)}$ α y β tal que

$$f(\beta)=f(\alpha)+\frac{f'(\alpha)}{1!}(\beta-\alpha)+\frac{f''(\alpha)}{2!}(\beta-\alpha)^2+\cdots+\frac{f^{(n-1)}(\alpha)}{(n-1)!}(\beta-\alpha)^{n-1}+\frac{f^{(n)}(\gamma)}{n!}(\beta-\alpha)^n.$$

DEMOSTRACION. Sea P el número real definido por medio de la relación

$$(28.1)\quad \frac{(\beta-\alpha)^n}{n!}P=f(\beta)-\left\{f(\alpha)+\frac{f'(\alpha)}{1!}(\beta-\alpha)+\cdots+\frac{f^{(n-1)}(\alpha)}{(n-1)!}(\beta-\alpha)^{n-1}\right\}.$$

y considere la función φ definida en J por

$$\varphi(x)=f(\beta)-\left\{f(x)+\frac{f'(x)}{1!}(\beta-x)+\cdots+\frac{f^{(n-1)}(x)}{(n-1)!}(\beta-x)^{n-1}+\frac{P}{n!}(\beta-x)^n\right\}.$$

Es claro que φ es continua en J y tiene una derivada en (a,b). Es evidente que $\varphi(\beta)=0$ y de la definición de P se deduce que $\varphi(\alpha)=0$. Por el teorema de Rolle, existe un punto γ entre α y β tal que $\varphi'(\gamma)=0$. Al calcular la derivada φ' (usando la fórmula común para la derivada de una suma y el producto de dos funciones), se obtiene la suma

$$\varphi'(x)=-\left\{f'(x)-f'(x)+\frac{f''(x)}{1!}(\beta-x)+\cdots+(-1)\frac{f^{(n-1)}(x)}{(n-2)!}(\beta-x)^{n-2}+\frac{f^{(n)}(x)}{(n-1)!}(\beta-x)^{n-1}-\frac{P}{(n-1)!}(\beta-x)^{n-1}\right\}$$
$$=\frac{P-f^{(n)}(x)}{(n-1)!}(\beta-x)^{n-1}.$$

Dado que $\varphi'(\gamma)=0$, entonces $P=f^{(n)}(\gamma)$, demostrándose la afirmación.

Q.E.D.

OBSERVACION. Al término residual

$$(28.2)\qquad R_n=\frac{f^{(n)}(\gamma)}{n!}(\beta-\alpha)^n$$

que se dio antes, con frecuencia se le llama la **forma de Lagrange** del residuo. Existen muchas otras expresiones para el residuo, pero por el momento sólo se mencionará la **forma de Cauchy** que asegura que para algún número θ con $0<\theta<1$,

$$(28.3)\qquad R_n=(1-\theta)^{n-1}\frac{f^{(n)}((1-\theta)\alpha+\theta\beta)}{(n-1)!}(\beta-\alpha)^n.$$

Para probar esta forma se puede repetir lo que se hizo hace un momento, excepto que del lado izquierdo de la ecuación (28.1) se pone $(\beta - x)Q/(n-1)!$ y en el último término es $(\beta - \alpha)Q/(n-1)!$ Los detalles quedan como ejercicio. (En la sección 31 se obtendrá otra forma en la que se hace uso de la integral para calcular el término residual.)

Ejercicios

28.A. Usando las fórmulas que aparecen en 28.1, probar que si $n = 0, 1, 2, \ldots,$ entonces las raíces de las funciones de Bessel J_n y J_{n+1} en $(0, +\infty)$ se entrelazan entre sí.

28.B. Probar que si $x > 0$, entonces

$$1 + \frac{x}{2} - \frac{x^2}{8} \leq \sqrt{1+x} \leq 1 + \frac{x}{2}.$$

28.C. Calcular $\sqrt{1.2}$ y $\sqrt{2}$. ¿Cual es la mejor aproximación de la que se puede estar seguro?

28.D. Obtener cálculos análogos a los del ejercicio 28.B para $(1+x)^{1/3}$ en el intervalo $[0, 7]$. Usar éstos para calcular $\sqrt[3]{1.5}$ y $\sqrt[3]{2}$.

28.E. Suponga que $0 < r < 1$ y $-1 < x$. Probar que se tiene $(1+x)^r \geq 1 + rx$ y que la igualdad es válida si y sólo si $x = 0$.

28.F. Se dice que una raíz x_0 de un polinomio p es **simple** (o que tiene **multiplicidad uno**) si $p'(x_0) \neq 0$, y que tiene **multiplicidad** n si $p(x_0) = p'(x_0) = \cdots = p^{(n-1)}(x_0) = 0$, pero $p^{(n)}(x_0) \neq 0$.

Si $a < b$ son raíces consecutivas de un polinomio, entonces existe un número impar (contando multiplicidades) de raíces de su derivada en (a, b).

28.G. Demostrar que si las raíces del polinomio p son todas reales, entonces las raíces de p' son todas reales. Si además las raíces de p son todas simples, entonces las raíces de p' son todas simples.

28.H. Si $f(x) = (x^2 - 1)^n$ y si p es la n-ésima derivada de f, entonces p es un polinomio de grado n cuyas raíces son simples y se encuentran en el intervalo abierto $(-1, 1)$.

28.I. Probar la forma de Cauchy del término residual R_n del teorema de Taylor dada en la fórmula (28.3).

28.J Una demostración para el teorema de Taylor 28.6 usando el teorema del valor medio de Cauchy se puede dar si

$$R(x) = f(x) - \left[f(\alpha) + \frac{x-\alpha}{1!} f'(\alpha) + \cdots + \frac{(x-\alpha)^{n-1}}{(n-1)!} f^{(n-1)}(\alpha)\right].$$

Probar que $R(\alpha) = R'(\alpha) = \cdots = R^{(n-1)}(\alpha) = 0$ y $R^{(n)}(x) = f^{(n)}(x)$. Obsérvese que existe γ_1 entre α y β tal que

$$\frac{R(\beta)}{(\beta-\alpha)^n} = \frac{R(\beta) - R(\alpha)}{(\beta-\alpha)^n - 0^n} = \frac{R'(\gamma_1)}{n(\gamma_1 - \alpha)^{n-1}}.$$

Continuar esto hasta encontrar que $R(\beta) = (\beta - \alpha)^n f^{(n)}(\gamma_n)/n!$ para alguna γ_n entre α y β.

28.K. Si $f(x) = e^x$, probar que el término residual en el teorema de Taylor converge a cero conforme $n \to \infty$ para cada una de las α, β. fijas.

28.L. Si $f(x) = \text{sen}\, x$, demostrar que el término residual del teorema de Taylor converge a cero conforme $n \to \infty$ para cada una de las α, β fijas.

28.M. Si $f(x) = (1+x)^m$ en donde $m \in \boldsymbol{Q}$, $|x| < 1$, las fórmulas comunes de diferenciación del cálculo y del teorema de Taylor dan la expresión

$$(1+x)^m = 1 + \binom{m}{1}x + \binom{m}{2}x^2 + \cdots + \binom{m}{n-1}x^{n-1} + R_n,$$

en donde R_n se puede expresar en la forma de Lagrange como $R_n = x^n f^{(n)}(\theta_n x)/n!$ en donde $0 < \theta_n < 1$. Demostrar que si $0 \leq x < 1$, entonces $\lim (R_n) = 0$. Probar que si $-1 < x < 0$, entonces no se puede usar el mismo argumento para probar que $(R_n) = 0$.

28.N. En el ejercicio anterior, usar la forma residual de Cauchy para obtener

$$R_n = \frac{m(m-1)\cdots(m-n+1)}{1\cdot 2\cdots(m-1)} \frac{(1-\theta_n)^{n-1}x^n}{(1+\theta_n x)^{n-m}},$$

en donde $0 < \theta_n < 1$. Cuando $|x| < 1$ probar que $|(1-\theta_n)/(1+\theta_n x)| < 1$, y demostrar que $\lim (R_n) = 0$.

28.O. Sea $f: \boldsymbol{R} \to \boldsymbol{R}$, si $f'(x)$ existe para $x \in \boldsymbol{R}$, y si $f''(a)$ existe, probar que

$$f''(a) = \lim_{h \to 0} \frac{f(a+h) - 2f(a) + f(a-h)}{h^2}.$$

Dar un ejemplo en que este límite exista pero en que la función no tenga una segunda derivada en a.

28.P. Sea $f_n(x) = |x|^{1+1/n}$ para x en $[-1, 1]$. probar que a cada f_n es diferenciable en $[-1, 1]$ y que (f_n) converge uniformemente en $[-1, 1]$ to $f(x) = |x|$.

Proyectos

28.α. En este proyecto se considera a la función exponencial desde el punto de vista del cálculo diferencial.

(a) Suponga ue una función E de $J = (a, b)$ a $\boldsymbol{R}$ tiene una derivada en todo punto de J y que $E'(x) = E(x)$ para toda $x \in J$. Observe que E tiene derivadas de todos los órdenes en J y todas son igual a E.

(b) Si $E(\alpha) = 0$ para alguna $\alpha \in J$, aplicar el teorema de Taylor 28.6 y el ejercicio 14.L para probar que $E(x) = 0$ para toda $x \in J$.

(c) Demostrar que existe cuando más una función E de $\boldsymbol{R}$ a $\boldsymbol{R}$ que satisface

$$E'(x) = E(x) \quad \text{para } x \in \boldsymbol{R}, \qquad E(0) = 1.$$

(d) Probar que si E satisface las condiciones del inciso *(c)* entonces también satisface la ecuación funcional

$$E(x+y) = E(x)E(y) \quad \text{para } x, y \in \boldsymbol{R}.$$

(Sugerencia: si $f(x) = E(x+y)/E(y)$, entonces $f'(x) = f(x)$ y $f(0) = 1$.)

(e) Sea (E_n) la sucesión de funciones definidas en $\boldsymbol{R}$ por medio de

$$E_1(x) = 1 + x, \qquad E_n(x) = E_{n-1}(x) + x^n/n!.$$

Sea A cualquier número positivo; si $|x| \leq A$ y si $m \geq n > 2A$, entonces

$$|E_m(x) - E_n(x)| \leq \frac{A^{n+1}}{(n+1)!}\left[1 + \frac{A}{n} + \cdots + \left(\frac{A}{n}\right)^{m-n}\right] < \frac{2A^{n+1}}{(n+1)!}.$$

Por lo tanto, la sucesión (E_n)converge uniformemente para $|x| \leq A$.

(f) Si (E_n) es la sucesión de funciones definidas en el inciso *(e)*, entonces

$$E_n'(x) = E_{n-1}(x), \qquad \text{para } x \in \boldsymbol{R}.$$

Probar que la sucesión (E_n)converge en $\boldsymbol{R}$ a una función E con las propiedades expuestas en el inciso *(c)*. Por lo tanto, E es la única función con estas propiedades.

(g) Sea E la función con $E' = E$ y $E(0) = 1$. Si e se define como el número

$$e = E(1),$$

entonces e está entre $2\frac{2}{3}$ y $2\frac{3}{4}$. *(Sugerencia:* $1 + 1 + \frac{1}{2} + \frac{1}{6} < e < 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{12}$. Más aún, se puede probar que $2.708 < 2 + \frac{17}{24} < e < 2 + \frac{13}{18} < 2.723$.)

28.β. En este proyecto se pueden usar los resultados del anterior. Sea E la única función en $\boldsymbol{R}$ tal que

$$E' = E \qquad \text{y} \qquad E(0) = 1$$

y sea $e = E(1)$.

(a) Probar que E es estrictamente creciente y que tiene rango $P = \{x \in \boldsymbol{R} : x > 0\}$.

(b) Sea L la función inversa de E, de tal manera que el dominio de L es P y su rango es todo $\boldsymbol{R}$. Demostrar que L es estrictamente creciente en P, que $L(1) = 0$, y que $L(e) = 1$.

(c) Probar que $L(xy) = L(x) + L(y)$ para todas x, y en P.

(d) Si $0 < x < y$, entonces

$$\frac{1}{y}(y - x) < L(y) - L(x) < \frac{1}{x}(y - x).$$

(Sugerencia: aplicar el teorema del valor medio a E.)

(e) La función L tiene una derivada para $x > 0$ y $L'(x) = 1/x$.

(f) El número e satisface

$$e = \lim\left(\left(1 + \frac{1}{n}\right)^n\right).$$

(Sugerencia: calcular $L'(1)$ usando la sucesión $((1 + 1/n))$ y la continuidad de E.)

28.γ. En este proyecto se introducirán el seno y el coseno.

(a) Defina a h en un intervalo $J = (a, b)$ a $\boldsymbol{R}$ tal que satisfaga

$$h''(x) + h(x) = 0$$

Para toda x en J. Demostrar que h tiene derivadas de todos los órdenes y que si hay un punto α en J tal que $h(\alpha) = 0$, $h'(\alpha) = 0$, entonces $h(x) = 0$ para toda $x \in J$. *(Sugerencia:* usar el teorema de Taylor 28.6.)

(b) Probar que existe cuando más una función C en $\boldsymbol{R}$ que satisface las condiciones

$$C'' + C = 0, \qquad C(0) = 1, \qquad C'(0) = 0,$$

y cuando más una función S en $\boldsymbol{R}$ que satisface

$$S'' + S = 0, \qquad S(0) = 0, \qquad S'(0) = 1.$$

(c) Defina una sucesión (C_n) por medio de

$$C_1(x) = 1 - x^2/2, \qquad C_n(x) = C_{n-1}(x) + (-1)^n \frac{x^{2n}}{(2n)!}.$$

Sea A cualquier número positivo; si $|x| \le A$ y si $m \ge n > A$, entonces

$$|C_m(x) - C_n(x)| \le \frac{A^{2n+2}}{(2n+2)!}\left[1 + \left(\frac{A}{2n}\right)^2 + \cdots + \left(\frac{A}{2n}\right)^{2m-2n}\right]$$
$$< \left(\frac{4}{3}\right)\frac{A^{2n+2}}{(2n+2)!}.$$

Por lo tanto, la sucesión (C_n) converge uniformemente para $|x| \le A$. Demostrar también que $C_n'' = -C_{n-1}$, y que $C_n(0) = 1$ y $C_n'(0) = 0$. Demostrar que el límite C de la sucesión (C_n) es la única función con las propiedades del inciso *(b)*.

(d) Defina a (S_n) por medio de

$$S_1(x) = x, \qquad S_n(x) = S_{n-1}(x) + (-1)^{n-1} \frac{x^{2n-1}}{(2n-1)!}.$$

Probar que (S_n) converge uniformemente para $|x| \le A$ a la función única S con las propiedades del inciso *(b)*.

(e) Demostrar que $S' = C$ y $C' = -S$.

(f) Probar la igualdad de Pitágoras $S^2 + C^2 = 1$. *(Sugerencia:* calcular la derivada de $S^2 + C^2$.)

28.δ. En este proyecto se continúa el análisis de las funciones seno y coseno. Se pueden usar libremente las propiedades demostradas en el proyecto anterior.

(a) Suponga que h es una función en $\boldsymbol{R}$ que satisface la ecuación

$$h'' + h = 0.$$

Probar que existen constantes α, β tales que $h = \alpha C + \beta S$. *(Sugerencia:* $\alpha = h(0)$, $\beta = h'(0)$.)

(b) La función C es par y S es impar en el sentido de que

$$C(-x) = C(x) \qquad \text{y} \qquad S(-x) = -S(x) \qquad \text{para toda } x \text{ en } \boldsymbol{R}$$

(c) Probar que las "fórmulas de suma"

$$C(x+y) = C(x)C(y) - S(x)S(y),$$
$$S(x+y) = S(x)C(y) + C(x)S(y),$$

son válidas para todas x, y en $\boldsymbol{R}$. *(Sugerencia:* sea y fija, defina $h(x) = C(x+y)$, y demuestre que $h'' + h = 0$.)

(d) Probar que las fórmulas de duplicación.

$$C(2x) = 2[C(x)]^2 - 1 = 2[S(x)]^2 + 1,$$
$$S(2x) = 2S(x)C(x),$$

son válidas para toda x en $\boldsymbol{R}$.

(e) Demostrar que C satisface la desigualdad

$$C_1(x) = 1 - \frac{x^2}{2} \leq C(x) \leq 1 - \frac{x^2}{2} + \frac{x^4}{24} = C_2(x).$$

Por lo tanto, la mínima raíz positiva γ de C está entre la raíz positiva de $x^2 - 2 = 0$ y la mínima raíz positiva de $x^4 - 12x^2 + 24 = 0$. Usando esto, demostrar que $\sqrt{2} < \gamma < \sqrt{3}$.

(f) Se define a π como la mínima raíz positiva de S. Probar que $\pi = 2\gamma$ y por lo tanto, que $2\sqrt{2} < \pi < 2\sqrt{3}$.

(g) Demostrar que tanto C como S son funciones periódicas con periodo 2π en el sentido de que $C(x+2\pi) = C(x)$ y $S(x+2\pi) = S(x)$ para toda x en $\boldsymbol{R}$. Demostrar también que

$$S(x) = C\left(\frac{\pi}{2} - x\right) = -C\left(x + \frac{\pi}{2}\right),$$

$$C(x) = S\left(\frac{\pi}{2} - x\right) = S\left(x + \frac{\pi}{2}\right),$$

para toda x en $\boldsymbol{R}$.

28.ε. Siguiendo el modelo de los dos proyectos anteriores, introducir el coseno y seno hiperbólico como funciones que satisfagan

$$c'' = c, \qquad c(0) = 1, \qquad c'(0) = 0,$$

$$s'' = s, \qquad s(0) = 0, \qquad s'(0) = 1,$$

respectivamente. Probar la existencia y la unicidad de estas funciones y demostrar que

$$c^2 - s^2 = 1.$$

Demostrar resultados análogos a (a)–(d) del proyecto 28.δ y probar que si la función exponencial se denota por medio de E entonces

$$c(x) = \tfrac{1}{2}(E(x) + E(-x)), \qquad s(x) = \tfrac{1}{2}(E(x) - E(-x)).$$

28.ζ. Una función φ en un intervalo J de $\boldsymbol{R}$ a $\boldsymbol{R}$ se dice que es **convexa (de punto medio)** siempre que

$$\varphi\left(\frac{x+y}{2}\right) \leq \tfrac{1}{2}(\varphi(x) + \varphi(y))$$

para cada x, y en J. (Geométricamente: el punto medio de cualquier cuerda de la curva $y = \varphi(x)$ está arriba o en la curva.) En este proyecto se habrá de suponer que φ es una función continua convexa.

(a) Si $n = 2^m$ y si $x_1, \ldots, x_n$ pertenecen a J, entonces

$$\varphi\left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right) \leq \frac{1}{n}(\varphi(x_1) + \cdots + \varphi(x_n)).$$

(b) Si $n < 2^m$ y si $x_1, \ldots, x_n$ pertenecen a J, sea x_j para $j = n+1, \ldots, 2^m$ igual a

$$\bar{x} = \left(\frac{x_1 + x_2 + \cdots + x_n}{n}\right).$$

Probar que se cumple la misma desigualdad que en el inciso *(a)*.

(c) Dado que φ es continua, probar que si x, y pertenecen a J y $t \in \boldsymbol{I}$, entonces

$$\varphi((1-t)x + ty) \leq (1-t)\varphi(x) + t\varphi(y).$$

(Geométricamente: toda la cuerda está arriba o en la curva.)

(d) Suponga que φ tiene una segunda derivada en J. Entonces, una condición necesaria y suficiente para que φ sea convexa en J es que $\varphi''(x) \geq 0$ para $x \in J$. *(Sugerencia:* para probar la condición necesaria, usar el ejercicio 28.0. Para probar la condición suficiente, usar el teorema de Taylor y desarrollar $\bar{x} = (x+y)/2$.)

(e) Si φ es una función convexa continua en J y si $x \leq y \leq z$ pertenecen a J, probar que

$$\frac{\varphi(y)-\varphi(x)}{y-x} \leq \frac{\varphi(z)-\varphi(x)}{z-x}.$$

Por lo tanto, si $w \leq x \leq y \leq z$ pertenecen a J, entonces

$$\frac{\varphi(x)-\varphi(w)}{x-w} \leq \frac{\varphi(z)-\varphi(y)}{z-y}.$$

(f) Demostrar que una función convexa continua φ en J tiene una derivada por la izquierda y una derivada por la derecha en todo punto. Más aún, el subconjunto en donde φ' no existe en numerable.

Sección 29 La integral de Riemann-Stieltjes

En esta sección se definirá la integral de Riemann-Stieltjes† de funciones acotadas en un intervalo compacto de $\boldsymbol{R}$. Como se supone que el lector está familiarizado, al menos informalmente, con la integral por algún curso de cálculo, no se hará muy extenso este tema.

El lector que continúe sus estudios de análisis matemático deseará familiarizarse pronto con la integral de Lebesgue que es más general. Sin embargo, dado que las integrales de Riemann y Riemann-Stieltjes son apropiadas para muchos propósitos y resultan más familiares al lector, es preferible tratarlas aquí y dejar para cursos posteriores la teoría de Lebesgue que es más avanzada.

Se habrán de considerar funciones acotadas de valor real en intervalos cerrados del sistema de números reales, se definirá la integral de una de estas funciones con respecto a otra y se obtendrán las propiedades principales de esta integral. El tipo de integración que se considera aquí es, en cierto sentido, más general que el que se ve en cursos anteriores y la generalidad que se

† (GEORGE FRIEDRICH) BERNHARD RIEMANN (1826-1866) fue hijo de un ministro campesino humilde y nació cerca de Hanover. Estudió en Gottingen y Berlín e impartió clases en Gottingen. Fue uno de los fundadores de la teoría de funciones analíticas pero también hizo contribuciones fundamentales a la geometría, a la teoría de números y a la física matemática.

THOMAS JOANNES STIELTJES (1856-1894) fue astrónomo y matemático alemán. Estudió en París con Hermite y obtuvo una cátedra en Toulouse. Su trabajo más famoso fue un informe sobre fracciones continuas, el problema del momento y la integral de Stieltjes, que se publicó en el ultimo año de su corta vida.

agrega lo hace muy útil en ciertas aplicaciones, especialmente en estadística. Al mismo tiempo, es muy poca la complicación extra, en la maquinaria teórica que requiere un análisis riguroso de la integral de Riemann ordinaria. Por lo tanto, vale la pena desarrollar este tipo de teoría de integración tanto como sus aplicaciones más frecuentes lo requieren.

Sean f y g funciones de valor real definidas en un intervalo cerrado $J = [a, b]$ de la recta real. *Se habrá de suponer que tanto f como g están acotadas en J;* esta hipótesis constante no se repetirá. Una **partición** de J es una colección finita de intervalos no traslapados cuya unión es J. Por lo general, una partición P se describe haciendo específico un conjunto finito de números reales $(x_0, x_1, \ldots, x_n)$ tal que

$$a = x_0 \leq x_1 \leq \cdots \leq x_n = b$$

y tal que los subintervalos que ocurren en la partición P son los intervalos $[x_{k-1}, x_k]$, $k = 1, 2, \ldots, n$. Más apropiadamente se hará referencia a los puntos terminales x_k, $k = 0, 1, \ldots, n$ como los **puntos de la partición** correspondiente a P. Sin embargo, en la practica a menudo es conveniente y no causa ninguna confusión, usar la palabra "partición" para designar la colección de subintervalos o bien la colección de puntos terminales de estos subintervalos. Por lo tanto, se escribe $P = (x_0, x_1, \ldots, x_n)$.

Si P y Q son particiones de J, se dice que Q es un **refinamiento de** P o que Q es **más fino que** P siempre que todo subintervalo en Q esté contenido en algún subintervalo en P. Esto es equivalente a la condición de que todo punto de la partición en P sea también un punto de la partición en Q. Por esta razón se escribe $P \subseteq Q$ cuando Q es un refinamiento de P.

29.1 DEFINICION. Si P es una partición de J, entonces una suma de Riemann-Stieltjes de f con respecto a g y correspondiente a $P = (x_0, x_1, \ldots, x_n)$ es un número real $S(P; f, g)$ de la forma

$$S(P; f, g) = \sum_{k=1}^{n} f(\xi_k)\{g(x_k) - g(x_{k-1})\}. \tag{29.1}$$

Se han elegido aquí número ξ_k que satisfacen

$$x_{k-1} \leq \xi_k \leq x_k \qquad \text{para} \qquad k = 1, 2, \ldots, n.$$

Observe que si la función g está dada por $g(x) = x$, entonces la expresión en la ecuación (29.1) se reduce a (29.2)

$$\sum_{k=1}^{n} f(\xi_k)(x_k - x_{k-1}). \tag{29.2}$$

A la suma (29.2) por lo general se la llama **suma de Riemann** de f correspondiente a la participación P y se puede interpretar como el área de la unión de rectángulos con lados $[x_{k-1}, x_k]$ y alturas $f(\xi_k)$. (figura 29.1) De manera que si la partición P es muy fina, se espera que la suma de Riemann (29.2) dé una aproximación al "área bajo la gráfica de f". Para una función general g, el lector deberá interpretar la suma de Riemann-Stieltjes (29.1) análoga a la suma de Riemann (29.2); excepto que en vez de

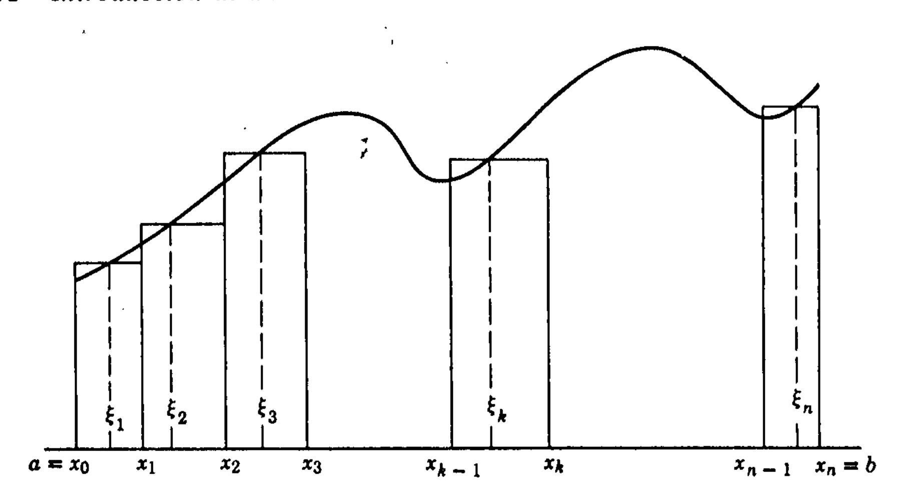


Figura 29.1 La suma de Riemann como un área.

considerar la longitud $x_k - x_{k-1}$ del subintervalo $[x_{k-1}, x_k]$, se está considerando otra medida de magnitud de este subintervalo, específicamente la diferencia $g(x_k) - g(x_{k-1})$. De modo que si $g(x)$ es la "masa" o "carga" total en el intervalo $[a, x]$, entonces $g(x_k) - g(x_{k-1})$designa la "masa" o "carga" en el subintervalo $[x_{k-1}, x_k]$. La idea es que se desea poder considerar medidas de magnitud de un intervalo distintas de longitud, de modo que se toman en cuenta las sumas (29.1) ligeramente más generales.

Se observará que ambas sumas (29.1) y (29.2) dependen de la elección de los "puntos intermedios" es decir, de los números ξ_k, $1 \leq k \leq n$. De modo que puede ser aconsejable introducir una notación que muestre la elección de estos números. Sin embargo, al introducir una partición más fina siempre se puede suponer que los puntos intermedios ξ_k son puntos de pertición. De hecho, si se introduce la partición $Q = (x_0, \xi_1, x_1, \xi_2, \ldots, \xi_n, x_n)$ y la suma $S(Q; f, g)$ en que los puntos intermedios se han tomado alternativamente de los puntos terminales a la derecha y a la izquierda en el subintervalo, entonces, la suma $S(Q; f, g)$da el mismo valor que la suma de (29.1).Se podría suponer siempre que la partición divide al intervalo en un número par de subintervalos y que los puntos intermedios son, alternativamente, los puntos terminales de la derecha y de la izquierda de estos subintervalos. Sin embargo, no será necesario requerir de este proceso de partición "modelo" ni tampoco será necesario exhibir estos puntos intermedios.

29.2. DEFINICION Se dice que f es **integrable** con respecto a g en J si existe un número real I tal que para todo número $\varepsilon > 0$ exista una partición P_ε de J tal que si P es cualquier refinamiento de P_ε y $S(P;f;g)$ es cualquier suma de Riemann-Stieltjes correspondientes a P, entonces

(29.3) $$|S(P; f, g) - I| < \varepsilon.$$

En este caso el número I esta determinado de manera única y se define por medio de

$$I = \int_a^b f\,dg = \int_a^b f(t)\,dg(t);$$

Se llama la **integral de Riemann-Stieltjes** de f con respecto a g sobre $J = [a, b]$. A la función f se la llama el **integrando** y a g el **integrante.** Algunas veces se dice que f es g-**integrable** si f es integrable con respecto a g. En el caso especial en que $g(x) = x$, si f es integrable con respecto a g,,generalmente se dice que f es **Riemann integrable.**

Antes de desarrollar cualquiera de las propiedades de la integral de Riemann-Stieltjes se verán algunos ejemplos. Con el objeto de que los cálculos sean sencillos, algunos de estos ejemplos se eligieron como casos extremos; ejemplos más típicos se encontrarán al combinar los que se dan en seguida.

29.3 EJEMPLOS. *(a)* Ya se ha visto que si $g(x) = x$, en entonces la integral se reduce a la integral común de Riemann Riemann del cálculo elemental.

(b) Si g es constante en el intervalo $[a, b]$, entonces cualquier función f es integrable con respecto a g y el valor de la integral es 0.

(c) Defínase g en $J = [a, b]$ como

$$g(x) = 0, \qquad x = a,$$
$$= 1, \qquad a < x \leq b.$$

Queda como ejercicio demostrar que una función f es integrable con respecto a g si y sólo si f es continua en a y que en este caso el valor de la integral es $f(a)$

(d) Sea c un punto interior del intervalo $J = [a, b]$ y defina g como

$$g(x) = 0, \qquad a \leq x \leq c,$$
$$= 1, \qquad c < x \leq b.$$

Demostrar como ejercicio que una función f es integrable con respecto a g si y sólo si es **continua** en c por la derecha (en el sentido de que para toda $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si $c \leq x < c + \delta(\varepsilon)$ y $x \in J$, entonces, se tiene $|f(x) - f(c)| < \varepsilon$). Si f satisface esta condición entonces el valor de la integral es $f(c)$. (Observe que la función integrante g es continua en c por la izquierda.)

(e) Modificando el ejemplo anterior, sea h la función definida por

$$h(x) = 0, \qquad a \leq x < c,$$
$$= 1, \qquad c \leq x \leq b.$$

Entonces, h es continua en c por la derecha y una función f es integrable con respecto a h si y sólo si f es continua en c por la izquierda. En este caso el valor de la integral es $f(c)$.

(f) Sean $c_1 < c_2$ puntos interiores de $J = [a, b]$ y defina g como

$$g(x) = \alpha_1, \qquad a \leq x \leq c_1,$$
$$= \alpha_2, \qquad c_1 < x \leq c_2,$$
$$= \alpha_3, \qquad c_2 < x \leq b.$$

Si f es continua en los puntos c_1, c_2, entonces f es integrable con respecto a g y

$$\int_a^b f\,dg = (\alpha_2 - \alpha_1)f(c_1) + (\alpha_3 - \alpha_2)f(c_2).$$

Tomando más puntos se puede obtener una suma en relación con los valores de f en puntos de J sopesados por los valores de los saltos de g en estos puntos.

(g) Sea f la función discontinua de Dirichlet (cf. ejemplo 20.5 *(g)*) definida por

$$f(x) = 1, \quad \text{si } x \text{ es racional}$$
$$= 0, \quad \text{si } x \text{ es irracional}$$

y sea $g(x) = x$. Considere estas fucniones en $I = [0, 1]$. Si una partición P consta de n subintervalos iguales, entonces eligiendo a k de los puntos intermedios en la suma $S(P; f, g)$ como racionales y el resto como irracionales $S(P; f, g) = k/n$. Se deduce que f no es Riemann integrable.

(h) Sea f la función definida en I por $f(0) = 1$, $f(x) = 0$ para irracional y $f(m/n) = 1/n$ cuando m y n son números naturales que no tienen factores comunes excepto 1. En el ejemplo 20.5 *(h)* se vio que f es continua en todo número irracional y discontinuo en todo número racional. Si $g(x) = x$, entonces queda como ejercicio demostrar que f es integrable con respecto a g y que el valor de la integral es 0.

29.4 CRITERIO DE CAUCHY PARA INTEGRABILIDAD. *La función f es integrable con respecto a g sobre $J = [a, b]$ si y sólo si para cada número $\varepsilon > 0$ hay una partición de Q_ε y de J tal que si P y Q son refinamientos de Q_ε y si $S(P; f, g)$ y $S(Q; f, g)$ son cualesquiera sumas de Riemann-Stieltjes correspondientes, entonces*

$$|S(P; f, g) - S(Q; f, g)| < \varepsilon. \tag{29.4}$$

DEMOSTRACION. Si f es integrable, existe una partición P_ε tal que si P, Q son refinamientos de P_ε, entonces cualesquiera sumas de Riemann-Stieltjes correspondientes satisfacen $|S(P; f, g) - I| < \varepsilon/2$ y $|S(Q; f, g) - I| < \varepsilon/2$. Usando la desigualdad del triángulo se obtiene (29.4).

Inversamente, supóngase que el criterio se satisface. Para demostrar que f es integrable con respecto a g, se debe producir el valor de su integral y usar la definición 29.2. Sea Q_1 una partición de J tal que si P y Q son refinamientos de Q_n entonces $|S(P; f, g) - S(Q; f, g)| < 1$. Por inducción se elige Q_1, como un refinamiento de Q_{n-1} tal que si P y Q son refinamientos de Q_n, entonces

$$|S(P; f, g) - S(Q; f, g)| < 1/n. \tag{29.5}$$

Considere sucesión $(S(Q_n; f, g))$ de números reales obtenida de esta manera. Dado que Q_n es un refinamiento de Q_m cuando $n \geq m$, esta sucesión de sumas es una sucesión de Cauchy de números reales, independientemente de

cómo se elijan los puntos intermedios. Por el teorema 16.10, la sucesión converge a algún número real L. Por lo que si $\varepsilon > 0$, hay un entero N tal que $2/N < \varepsilon$ y

$$|S(Q_N; f, g) - L| < \varepsilon/2.$$

Si P es un refinamiento de Q_N, entonces de la construcción de Q_N se deduce que

$$|S(P; f, g) - S(Q_N; f, g)| < 1/N < \varepsilon/2.$$

Por lo tanto, para cualquier refinamiento P de Q_N y cualquier suma de Riemann-Stieltjes correspondiente se tiene

$$|S(P; f, g) - L| < \varepsilon, \tag{29.6}$$

Esto prueba que f es integrable con respecto a g sobre J y que el valor es esta integral es L. Q.E.D.

Algunas propiedades de la integral

La siguiente propiedad en ocasiones es conocida como la **bilinealidad** de la integral de Riemann-Stieltjes.

29.5 TEOREMA *(a) Si f_1, f_2 son integrables con respecto a g en J y α, β son números reales entonces $\alpha f_1 + \beta f_2$ es integrable con respecto a g en J y*

$$\int_a^b (\alpha f_1 + \beta f_2)\, dg = \alpha \int_a^b f_1\, dg + \beta \int_a^b f_2\, dg. \tag{29.7}$$

(b) Si f es integrable con a g_1 y g_2 en J y α, β son números reales, entonces f es integrable con respecto a $g = \alpha g_1 + \beta g_2$ en J y

$$\int_a^b f\, dg = \alpha \int_a^b f\, dg_1 + \beta \int_a^b f\, dg_2. \tag{29.8}$$

DEMOSTRACION *(a)* Sea $\varepsilon > 0$ y sean $P_1 = (x_0, x_1, \ldots, x_n)$ y $P_2 = (y_0, y_1, \ldots, y_m)$ particiones de $J = [a, b]$ tales que si Q es un refinamiento tanto de P_1 como de P_2, entonces para cualesquiera sumas de Riemann-Stieltjes correspondientes se tiene

$$|I_1 - S(Q; f_1, g)| < \varepsilon, \qquad |I_2 - S(Q; f_2, g)| < \varepsilon.$$

Sea P_ε una partición de J que sea un refinamiento de P_1 y de P_2 (por ejemplo, todos los puntos de partición en P_1 y P_2 se combinan para formar P_ε). Si Q es una partición de J tal que $P_\varepsilon \subseteq Q$, entonces las dos relaciones anteriores aún son válidas. Cuando se usan los mismos puntos intermedios, evidentemente se tiene

$$S(Q;\alpha f_1+\beta f_2, g)=\alpha S(Q;f_1, g)+\beta S(Q;f_2, g).$$

De esto y de las desigualdades anteriores se deduce que

$$|\alpha I_1+\beta I_2-S(Q;\alpha f_1+\beta f_2, g)|=|\alpha\{I_1-S(Q;f_1, g)\}+\beta\{I_2-S(Q;f_2, g)\}|$$
$$\leq(|\alpha|+|\beta|)\varepsilon.$$

Esto prueba que $\alpha I_1+\beta I_2$ es la integral $\alpha f_1+\beta f_2$ con respecto a g. Esto demuestra la parte *(a)*; la demostración de la parte *(b)* es análoga y se deja al lector. Q.E.D.

Existe otra útil propiedad de sumabilidad que tiene la integral de Riemann-Stieltjes; a saber, con respecto al intervalo sobre el cual se extiende la integral. Es con el objeto de obtener el siguiente resultado que se usó el tipo de límite introducido en la definición 29.2. Un tipo más restrictivo de límite sería requerir de la desigualdad (29.3), Para cualquier suma de Riemann-Stieltjes, que corresponda a una partición $P=(x_0, x_1, \ldots, x_n)$ tal que

$$\|P\|=\sup\{x_1-x_0, x_2-x_1, \ldots, x_n-x_{n-1}\}<\delta(\varepsilon).$$

Este tipo de límite por lo general se usa al definir la integral de Riemann y algunas veces al definir la integral de Riemann-Stieltjes. Sin embargo, muchos autores emplean la definición que se dio, debida a S Pollard, ya que amplía ligeramente la clase de funciones integrables. Como resultado de esta ampliación, el siguiente resultado es válido sin ninguna restricción adicional. Véanse los ejercicios 29.P-R.

29.6 TEOREMA. *(a) Suponga que* $a\leq c\leq b$ *y que f es integrable con respecto a g sobre ambos subintervalos* $[a, c]$ *y* $[c, b]$. *Entonces, f es integrable con respecto a g en el intervalo* $[a, b]$ *y*

$$\int_a^b f\,dg=\int_a^c f\,dg+\int_c^b f\,dg. \tag{29.9}$$

(b) Sea f integrable con respecto a g en el intervalo $[a, b]$ *y sea c tal que satisfaga* $a\leq c\leq b$. *Entonces f es integrable con respecto a g en los subintervalos* $[a, c]$ *y* $[c, b]$ *y la fórmula (29.9) es válida.*

DEMOSTRACION. (a) Si $\varepsilon>0$, sea P'_ε una partición de $[a, c]$ tal que si P' es un refinamiento de P'_ε, entonces la desigualdad (29.3) es válida para cualquier suma de Riemann-Stieltjes. Sea P''_ε una partición correspondiente de $[c, b]$. Si P_ε es la partición de $[a, b]$ formada al usar los puntos de partición en P'_ε y P''_ε, y si P es un refinamiento de P_ε, entonces

$$S(P;f, g)=S(P';f, g)+S(P'';f, g),$$

en donde P', P'' designan las particiones de $[a, c], [c, b]$ introducidas por P y en donde se usan los puntos intermedios correspondientes. Por lo tanto, se tiene

$$\left|\int_a^c f\,dg+\int_c^b f\,dg-S(P;f, g)\right|$$

$$\leq \left| \int_a^c f\,dg - S(P'; f, g) \right| + \left| \int_c^b f\,dg - S(P''; f, g) \right| < 2\varepsilon.$$

Se deduce que f es integrable con respecto a g sobre $[a, b]$ y que el valor de su integral es

$$\int_a^c f\,dg + \int_c^b f\,dg.$$

(b) Se usará el criterio de Cauchy 29.4 para demostrar que f es integrable sobre $[a, c]$. Dado que f es integrable sobre $[a, b]$, dada $\varepsilon > 0$ hay una partición Q_ε de $[a, b]$ tal que si P, Q son refinamientos de Q_ε, la relación (29.4) es válida para cualesquiera sumas de Riemann-Stieltjes correspondientes. Es claro que se puede suponer que el punto c pertenece a Q_ε y Q_ε' será la partición de $[a, c]$ que conste de aquellos puntos de Q_ε que pertenecen a $[a, c]$. Supóngase que P' y Q' son particiones de $[a, c]$ que son refinamientos de Q_ε' y extiéndanse a particiones P y Q de $[a, b]$ usando los puntos en Q_ε que pertenezcan a $[c, b]$. Dado que P, Q son refinamientos de Q_ε, la relación (29.4) es válida. Sin embargo a partir del hecho de que P, Q son idénticas en $[c, b]$ es claro que si se usan los mismos puntos intermedios entonces.

$$|S(P'; f, g) - S(Q'; f, g)| = |S(P; f, g) - S(Q; f, g)| < \varepsilon.$$

Por lo tanto, el criterio de Cauchy prueba la integrabilidad de f con respecto a g sobre el subintervalo $[a, c]$ un argumento análogo también es aplicable al intervalo $[c, b]$. Una vez conocida esta integrabilidad, la parte *(a)* proporciona la validez de la fórmula (29.9). Q.E.D.

Hasta este momento no se han intercambiado los papeles del integrando f y del integrante g y podría no haberle ocurrido al lector que sería posible hacerlo. Aun cuando el siguiente resultado no es exactamente lo mismo que la "fórmula de integración por partes" de cálculo, la relación es muy estrecha y algunas veces se conoce a este resultado por el mismo nombre.

29.7 INTEGRACION POR PARTES *Una función f es integrable con respecto a g sobre $[a, b]$ si y sólo si g es integrable con respecto a f sobre $[a, b]$. En este caso.*

$$\int_a^b f\,dg + \int_a^b g\,df = f(b)g(b) - f(a)g(a). \tag{29.10}$$

DEMOSTRACION. Se va a suponer que f es integrable con respecto a g. Sea $\varepsilon > 0$ y sea P_ε una partición de $[a, b]$ tal que si Q es un refinamiento de P_ε y $S(Q; f, g)$ es cualquier suma de Riemann-Stieltjes correspondiente, entonces

$$\left| S(Q; f, g) - \int_a^b f\,dg \right| < \varepsilon. \tag{29.11}$$

Ahora, sea p un refinamiento de P_ε y considere una suma de Riemann-Stieltjes $S(P; g, f)$ dada por

$$S(P; g, f) = \sum_{k=1}^{n} g(\xi_k)\{f(x_k) - f(x_{k-1})\},$$

en donde $x_{k-1} \leq \xi_k \leq x_k$. Sea $Q = (y_0, y_1, \ldots, y_{2n})$ la partición de $[a, b]$ obtenida usando a ξ_k y x_k como puntos de partición de donde $y_{2k} = x_k$ y $y_{2k-1} = \xi_k$. Sumar y restar los términos $f(y_{2k})g(y_{2k})$ para $k = 0, 1, \ldots, n$, $S(P; g, f)$ y ordenarlos de tal manera que

$$S(P; g, f) = f(b)g(b) - f(a)g(a) - \sum_{k=1}^{2n} f(\eta_k)\{g(y_k) - g(y_{k-1})\},$$

en donde los puntos intermedios η_k corresponden a los puntos x_j. De manera que se tiene

$$S(P; g, f) = f(b)g(b) - f(a)g(a) - S(Q; f, g),$$

en donde la partición $Q = (y_0, y_1, \ldots, y_{2n})$ es un refinamiento de P_ε. por la fórmula (9.11).

$$\left| S(P; g, f) - \left\{ f(b)g(b) - f(a)g(a) - \int_a^b f\, dg \right\} \right| < \varepsilon$$

siempre que p sea un refinamiento de P_ε. Esto demuestra que g es integrable con respecto a f sobre $[a, b]$ y se prueba la fórmula (29.10).

Q.E.D.

Modificación de la integral

Cuando la función integrante g tiene una derivada continua, es posible, y con frecuencia conveniente, reemplazar la integral de Riemann-Stieltjes por una integral de Riemann. En seguida se prueba la validez de esta reducción.

29.8 TEOREMA. *Si la derivada g' existe y es continua en J y si f es integrable con respecto a g entonces el producto fg' es Riemann integrable y*

(29.12) $$\int_a^b f\, dg = \int_a^b fg'.$$

DEMOSTRACION. La hipótesis implica que g' es uniformemente continua en J. Si $\varepsilon > 0$, sea $P = (x_0, x_1, \ldots, x_n)$ una partición de J tal que si ξ_k y ζ_k pertenecen a $[x_{k-1}, x_k]$ entonces $|g'(\xi_k) - g'(\zeta_k)| < \varepsilon$. Se considera la diferencia de la suma de Riemann-Stieltjes $S(P; f, g)$ y la suma de Riemann $S(P; fg')$, usando los mismos puntos intermedios ξ_k. Al hacer esto se tiene una suma de términos de la forma

$$f(\xi_k)\{g(x_k) - g(x_{k-1})\} - f(\xi_k)g'(\xi_k)\{x_k - x_{k-1}\}.$$

Si se aplica el teorema del valor medio 27.6 a g, la diferencia se puede escribir en la forma

$$f(\xi_k)\{g'(\zeta_k)-g'(\xi_k)\}(x_k-x_{k-1}),$$

en donde ζ_k es algún punto en el intervalo $[x_{k-1}, x_k]$. Dado que este término está dominado por $\varepsilon\,\|f\|\,(x_k-x_{k-1})$, se concluye que

$$|S(P;f,g)-S(P;fg')|\leq\varepsilon\,\|f\|\,(b-a),$$

siempre que la partición p sea lo suficientemente fina. Dado que la integral del lado izquierdo de (29.11) existe y es el límite de las sumas de Riemann-Stieltjes $S(P;f,g)$, se deduce que también existe la integral del lado derecho de (29.12) y que por lo tanto la igualdad es válida. Q.E.D.

Como extensión de este resultado, véase el teorema 30.13.

29.9 EJEMPLOS *(a)* De algunos resultados que se demostrarán en la sección 3O se deduce que $f(x)=x$ es integrable con respecto a $g(x)=x^2$ en $J=[0,1]$. Suponiendo esto, el teorema 29.8 muestra que

$$\int_0^1 x\,d(x^2)=\int_0^1 x\cdot 2x\,dx=\tfrac{2}{3}x^3\Big|_0^1=\frac{2}{3}.$$

(Aquí se han utilizado resultados de cálculo que se demostrarán en la sección 30.)

(b) Si se aplica el teorema 29.7 sobre integración por partes a las funciones de *(a)* se obtiene

$$\int_0^1 x\,d(x^2)=x^3\Big|_0^1-\int_0^1 x^2\,dx=1-\tfrac{1}{3}x\Big|_0^1=\frac{2}{3}.$$

(c) De algunos resultados que se demostrarán en la sección 30 se deduce que $f(x)=$ *sen* x es integrable con respecto a f en $J=[0,\pi/2]$. Soponiendo esto, se tiene

$$\int_0^{\pi/2}\operatorname{sen}x\,d(\operatorname{sen}x)=\int_0^{\pi/2}\operatorname{sen}x\cos x\,dx=\tfrac{1}{2}(\operatorname{sen}x)^2\Big|_0^{\pi/2}=\frac{1}{2}.$$

(d) Si se aplica el teorema 29.7 sobre integración por partes al inciso *(c)* se obtiene

$$\int_0^{\pi/2}\operatorname{sen}x\,d(\operatorname{sen}x)=(\operatorname{sen}x)^2\Big|_0^{\pi/2}-\int_0^{\pi/2}\operatorname{sen}x\,d(\operatorname{sen}x),$$

por lo que se deduce que

$$\int_0^{\pi/2}\operatorname{sen}x\,d(\operatorname{sen}x)=\tfrac{1}{2}(\operatorname{sen}x)^2\Big|_0^{\pi/2}=\frac{1}{2}.$$

(e) Se introduce la **función del entero máximo,** de $\boldsymbol{R}$ a $\boldsymbol{R}$, a la que se denota por medio del símbolo especial $[\,\cdot\,]$ y se define como la función tal que si

$x \in \boldsymbol{R}$, entonces $[x]$ es el entero mayor, menor o igual a x. De donde $[\pi]=3, [e]=2, [-2.5]=-3$. El lector deberá dibujar una gráfica de esta función y podrá observar que es continua por la derecha con saltos igual a 1 en los enteros. Se deduce que si f es continua en $[0, 5]$, entonces f es integrable con respecto a $g(x)=[x]$, $x \in [0, 5]$, y que

$$\int_0^5 f(x)\, d([x]) = \sum_{j=1}^{5} f(j).$$

(f) De algunos resultados de la sección 30 se deduce que $f(x)=x^2$ es integrable con respecto a $g_1(x)=x$ y a $g_2(x)=[x]$ en $[0, 5]$. Por lo tanto, es integrable con respecto a $g(x)=x+[x]$ y se tiene

$$\int_0^5 x^2\, d(x+[x]) = \int_0^5 x^2\, dx + \int_0^5 x^2\, d([x])$$
$$= \tfrac{1}{3}5^3 + 1^2 + 2^2 + 3^2 + 4^2 + 5^2.$$

Ejercicios

29.A. Si f es constante en el intervalo $[a, b]$, entonces es integrable con respecto a cualquier función g y

$$\int_a^b f\, dg = f(a)\{g(b)-g(a)\}.$$

29.B. Si g es como en el ejemplo 29.3 (c), demostrar que f es integrable a g si y sólo si f es continua en a.

29.C. Defina a g en $\boldsymbol{I}=[0, 1]$ como $g(x)=0$ para $0 \le x \le \frac{1}{2}$ y $g(x)=1$ para $\frac{1}{2} < x \le 1$. Probar que f es integrable con respecto a g en $\boldsymbol{I}$ si y sólo si es continua por la derecha en $\frac{1}{2}$ En este caso el valor de la integral es $f(\frac{1}{2})$.

29.D. Probar que la función f dada en el ejemplo 29.3 *(h)* es Riemann integrable en $\boldsymbol{I}$ y que el valor de su integral es 0.

29.E. Si f es integrable en $[a, b]$ con respecto a f, entonces

$$\int_a^b f\, df = \tfrac{1}{2}\{(f(b))^2 - (f(a))^2\}.$$

(a) Demostrar esto examinando las dos sumas de Riemann-Stieltjes para una partición $P=(x_0, x_1, \ldots, x_n)$ obtenida tomando $\xi_k = x_{k-1}$ y $\xi_k = x_k$.

(b) Demostrar esto usando el teorema de integración por partes 29.7.

29.F. Demostrar directamente que si f es la función del entero máximo $f(x)=[x]$ definida en el ejemolo 29.9 *(e)* entonces f *no* es integrable con respecto a f el intervalo $[0, 2]$.

29.G. Si f es Riemann integrable en $[0, 1]$, entonces

$$\int_0^1 f = \lim \left(\frac{1}{n}\sum_{k=1}^{n} f\left(\frac{k}{n}\right)\right).$$

29 H. Demostrar que si g *no* es integrable en $[0, 1]$, entonces la sucesión de promedios

$$\left(\frac{1}{n}\sum_{k=1}^{n} g\left(\frac{k}{n}\right)\right)$$

puede ser o puede no ser convergente.

29.I Demostrar que la función h, definida en I, por $h(x)=x$ para x racional y $h(x)=0$ para x irracional, no es Riemann integrable en I.

29.J. Suponga que f es Riemann integrable en $[a, b]$. Si f_1 es una función de $[a, b]$ a R tal que $f_1(x)=f(x)$ excepto para un *número finito* de puntos en $[a, b]$, demostrar que f_1 es Riemann integrable y que

$$\int_a^b f_1 = \int_a^b f.$$

(De modo que se puede cambiar el valor de una función Riemann integrable, o dejarla indefinida, en un número finito de puntos.)

2O.K. Dar un ejemplo para mostrar que la conclusión del ejercicio anterior puede no ser válida si el número de puntos excepcionales es infinito.

29.L. Sea $c \in (a, b)$ y defina a k en $[a, b]$ como $k(c)=1$ y $k(x)=0$ para $x \in [a, b]$, $x \neq c$. Si $f:[a, b] \to \mathbf{R}$ es continua en c, probar directamente que f es k-integrable, que k es f-integrable y que

$$\int_a^b f\,dk = \int_a^b k\,df = 0.$$

29.M. Suponga que f es g-integrable en $[a, b]$. Si $g_1:[a, b] \to \mathbf{R}$ es tal que $g_1(x)=g(x)$ excepto para un número finito de puntos en (a, b) en los que f es continua, entonces f es g_1-integrable y

$$\int_a^b f\,dg_1 = \int_a^b f\,dg.$$

29.N. Suponga que g es continua en $[a, b]$, que $x \mapsto g'(x)$ existe y es continua en $[a, b]\setminus\{c\}$, y que los límites unilaterales

$$g'(c-) = \lim_{\substack{x\to c\\ x<c}} g'(x), \qquad g'(c+) = \lim_{\substack{x\to c\\ x>c}} g'(x)$$

existen. Si f es integrable con respecto a g en $[a, b]$, entonces fg' se puede definir en c para ser Riemann-integrable en $[a, b]$ y tal que

$$\int_a^b f\,dg = \int_a^b fg'.$$

(*Sugerencia:* considere el ejercicio 27.N.)

29.O. Si f es Riemann-integrable en $[-5, 5]$, probar que f es integrable con respecto a $g(x)=|x|$ y

$$\int_{-5}^{5} f\,dg = \int_0^5 f - \int_{-5}^{0} f.$$

29.P. Si $P=(x_0, x_1, \ldots, x_n)$ es una partición de $J=[a, b]$, defina a $\|P\|$ como

$$\|P\| = \sup\{x_j - x_{j-1} : j=1, 2, \ldots, n\};$$

a $\|P\|$ se le llama la **norma** de partición P. Defina a f como(*)**-integrable** con respecto a g en J siempre que exista un número A con la propiedad de que si $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si $\|P\| < \delta(\varepsilon)$ y si $S(P; f, g)$ es cualquier suma de Riemann-Stieltjes correspondiente, entonces $|S(P; f, g) - A| < \varepsilon$. Si esto se cumple, al número A se le llama la (*)**-integral** de f con respecto a g en J. Demostrar que si f es (*)-integrable con respecto a g (en el sentido de la definición 29.2) y que los valores de estas integrales son iguales.

29.Q. Defina a g en I como en el ejercicio 29.C. Demostrar que una función acotada f es(*)-integrable con respecto a g, en el sentido del ejercicio anterior, si y sólo si f es continua en $\frac{1}{2}$ cuando el valor de la(*)-integral es $f(\frac{1}{2})$. Si h se define por medio de

$$h(x) = 0, \qquad 0 \leq x < \tfrac{1}{2},$$
$$= 1, \qquad \tfrac{1}{2} \leq x \leq 1,$$

entonces h es(*)-integrable con respecto a g en $[0, \frac{1}{2}]$ y en $[\frac{1}{2}, 1]$ pero no es(*)-integrable con respecto a g en $[0, 1]$. Por lo que el teorema 29.6(a) puede no cumplirse para la (*)-integral.

29.R. Sea $g(x) = x$ para $x \in J$. Demostrar que para este integrante una función f es integrable en el sentido de la definición 29.2 si y sólo si es(*)-integrable en el sentido del ejercicio 29.P.

29.S. Sea f Riemann integrable en J y sea $f(x) \geq 0$ para $x \in J$. Si f es continua en un punto $c \in J$ y si $f(c) > 0$, entonces.

$$\int_a^b f > 0.$$

29.T. Sea f Riemann integrable en J y sea $f(x) > 0$ para $x \in J$. Demostrar que

$$\int_a^b f > 0.$$

(*Sugerencia:* para cada $n \in \mathbf{N}$, sea H_n la cerradura del conjunto de puntos $x \in J$ tales que $f(x) > 1/n$ y aplicar el ejercicio 11.N.)

Proyectos

29.α. La siguiente esquematización se usa algunas veces como vía de acceso a la integral de Riemann-Stieltjes cuando la función integrante es monótonamente creciente en el intervalo J. (Este desarrollo tiene la ventaja de que usa la definición de las integrales superior e inferior que siempre existen para una función F. Sin embargo, tiene la desventaja de que le aumenta una restricción a g y tiende a dañar, en cierto sentido, la simetría de la integral de Riemann-Stieltjes dada por el teorema de integración por partes 29.7) Si $P = (x_0, x_1, \ldots, x_n)$ es una partición de $J = [a, b]$ y f es una función acotada en J, defina a m_j, M_j como el infimo y el supremo de $\{f(x) : x_{j-1} \leq x \leq x_j\}$, respectivamente. Correspondientes a la partición P, defina las **sumas inferior** y **superior** de f con respecto a g como

$$L(P; f, g) = \sum_{j=1}^{n} m_j\{g(x_j) - g(x_{j-1})\},$$

$$U(P; f, g) = \sum_{j=1}^{n} M_j\{g(x_j) - g(x_{j-1})\}.$$

(a) Si $S(P; f, g)$ es cualquier suma de Riemann-Stieltjes correspondiente a P, entonces

$$L(P; f, g) \leq S(P; f, g) \leq U(P; f, g).$$

(b) Si $\varepsilon > 0$ entonces existe una suma de Riemann-Stieltjes $S_1(P; f, g)$ correspondiente a P tal que

$$S_1(P; f, g) \leq L(P; f, g) + \varepsilon,$$

y existe una suma de Riemann-Stieltjes $S_2(P; f, g)$ correspondiente a P tal que

$$U(P; f, g) - \varepsilon \leq S_2(P; f, g).$$

(c) Si P y Q son particiones de J y si Q es un refinamiento de P (es decir, $P \subseteq Q$), entonces

$$L(P; f, g) \leq L(Q; f, g) \leq U(Q; f, g) \leq U(P; f, g).$$

(d) Si P_1 y P_2 son cualesquiera particiones de J, entonces $L(P_1; f, g) \leq U(P_2; f, g)$. *(Sugerencia:* sea Q una partición que es refinamiento tanto de P_1 como de P_2 y aplicar (c).]

(e) Defina las **integrales inferior** y **superior** de f con respecto a g, respectivamente, como

$$L(f, g) = \sup\{L(P; f, g)\},$$
$$U(f, g) = \inf\{U(P; f, g)\};$$

aquí, el supremo y el ínfimo se toman sobre todas las particiones P de J. Demostrar que $L(f, g) \leq U(f, g)$.

(f) Demostrar que f es integrable con respecto a la función creciente g si y sólo si las integrales inferior y superior introducidas en *(e)* son iguales. En este caso, el valor común de estas integrales es igual a

$$\int_a^b f\,dg.$$

Demostrar que f es integrable con respecto a g si y sólo si la siguiente condición de Riemann se cumple: para toda $\varepsilon > 0$ existe una partición P tal que $U(P; f, g) - L(P; f, g) < \varepsilon$.

(g) Si f_1 y f_2 están acotadas en J, entonces las integrales inferior y superior de $f_1 + f_2$ satisfacen

$$L(f_1 + f_2, g) \geq L(f_1, g) + L(f_2, g),$$
$$U(f_1 + f_2, g) \leq U(f_1, g) + U(f_2, g).$$

Probar que la desigualdad propia puede ser válida en estas relaciones.

29. β. Este y los dos siguientes proyectos introducen y analizan la importante clase de funciones que tienen "variación acotada" en un intervalo compacto. Sea $f : [a, b] \to \boldsymbol{R}$ dada; si $P = (a = x_0 < x_1 < \cdots < x_n = b)$ es una partición de $[a, b]$, defínase $v_f(P)$ como

i254 (1494x2426x2 pcx)

$$v_f(P) = \sum_{k=1}^{n'} |f(x_k) - f(x_{k-1})|.$$

Si el conjunto $\{v_f(P) : P$ es una partición de $[a, b]\}$ es acotado, se dice que f tiene **variación acotada** en $[a, b]$. La colección de todas las funciones que tienen variación acotada en $[a, b]$ se denota por $BV([a, b])$ o con $BV[a, b]$. Si $f \in BV[a, b]$, entonces se define

$$V_f[a, b] = \sup\{v_f(P) : P \text{ es una partición de } [a, b]\}.$$

Al número $V_f[a, b]$ se le llama la **variación total** de f en $[a, b]$. Demostrar que $V_f[a, b] = 0$ si y sólo si f es una función constante en $[a, b]$.

(a) Si $f : [a, b] \to \boldsymbol{R}$, si P y Q son particiones de $[a, b]$, y si $P \supseteq Q$, demostrar que $v_f(P) \geq v_f(Q)$. Si $f \in BV[a, b]$, demostrar que existe una sucesión (P_n) de particiones de $[a, b]$ tal que $V_f[a, b] = \lim(v_f(P_n))$.

(b) Si f es monótonamente creciente en $[a, b]$, demostrar que $f \in BV[a, b]$ y que $V_f[a, b] = f(b) - f(a)$. ¿Qué sucede si f es monótonamente decreciente en $[a, b]$?

(c) Si $g : [a, b] \to \boldsymbol{R}$ satisface la condición de Lipschitz $|g(x) - g(y)| \leq M|x - y|$ para todas x, y en $[a, b]$, demostrar que $g \in BV[a, b]$ y que $V_g[a, b] \leq M(b - a)$. Si $|h'(x)| \leq M$ para toda $x \in [a, b]$, entonces $h \in BV[a, b]$ y $V_h[a, b] \leq M(b - a)$. Más aún, considere $k(x) = \sqrt{x}$ en $[0, 1]$.

(d) Defina a $f : [0, 1] \to \boldsymbol{R}$ como $f(x) = 0$ para $x = 0$ y $f(x) = \operatorname{sen}(1/x)$ para $0 < x \leq 1$. Demostrar que f no tiene variación acotada en $[0, 1]$. Si g se define como $g(x) = xf(x)$ para $x \in [0, 1]$, demostrar que g es continua pero no tiene variación acotada en $[0, 1]$. Sin embargo, si h se define como $h(x) = x^2 f(x)$ para $x \in [0, 1]$, demostrar que h sí tiene variación acotada en $[0, 1]$.

(e) Si $f \in BV$ sí $[a, b]$, demostrar que $|f(x)| \leq |f(a)| + V_f[a, b]$ para toda $x \in [a, b]$, de tal manera que f está acotada en $I = [a, b]$ y $\|f\|_I \leq |f(a)| + V_f[a, b]$.

(f) Si $f, g \in BV[a, b]$ y $\alpha \in \mathbf{R}$, demostrar que αf y $f + g$ pertenecen a $BV[a, b]$ y que

$$V_{\alpha f}[a, b] = |\alpha|\, V_f[a, b],$$

$$V_{f+g}[a, b) \leq V_f[a, b] + V_g[a, b].$$

por lo tanto, $BV[a, b]$ es un espacio vectorial de funciones.

(g) Si $f, g \in BV[a, b]$, demostrar que el producto fg y que pertenece a $BV[a, b]$

$$V_{fg}[a, b] \leq \|f\|_I\, V_g[a, b] + \|g\|_I\, V_f[a, b].$$

Demostrar que el cociente de dos funciones en $BV[a, b]$ puede no pertenecer a $BV[a, b]$.

(h) Demostrar que la aplicación $f \mapsto V_f[a, b]$ no es una norma en el espacio vectorial $BV[a, b]$, pero la aplicación

$$f \mapsto \|f\|_{BV} = |f(a)| + V_f[a, b]$$

es una norma en este espacio.

29.γ. Se continúa con el estudio de funciones de variación acotada en un intervalo $[a, b] \subseteq \boldsymbol{R}$.

(a) Si $f \in BV[a, b]$ y si $c \in (a, b)$, demostrar que las restricciones de f a $[a, c]$ y $[c, b]$ tienen variación acotada en estos intervalos y que

$$V_f[a, b] = V_f[a, c] + V_f[c, b].$$

A la inversa, si $g:[a, b] \to \mathbf{R}$ es tal que para alguna $c \in (a, b)$ se tiene $g \in BV[a, c]$ y $g \in BV[a, b]$, entonces $g \in BV\,[a, b]$.

(b) Si xxxxxxxxxxxi entonces se define $p_f(x) = V_f[a, x]$ para $x \in (a, b]$, y $p_f(a) = 0$. Demostrar que p_f es una función creciente en $[a, b]$.

(c) Observe que si $a \le x \le y \le b$, entonces

$$f \in BV\,[a, b],\; f(y) - f(x) \le V_f[x, y].$$

Demostrar que si se define $n_f(x) = p_f(x) - f(x)$para $x \in [a, b]$, entonces n_f es una función creciente.

(d) Demostrar que una función $f:[a, b] \to \mathbf{R}$ pertenece a $BV[a, b]$ si y sólo si es la diferencia de dos funciones crecientes.

(e) Si $f \in BV[a, b]$ es continua por la derecha en un punto $c \in [a, b)$, y si $\varepsilon > 0$, demostrar que existe $\delta > 0$ una partición tal que si $Q = (c < x_1 < \cdots < x_n = b)$es una partición suficientemente fina de $[c, b]$ con $x_1 - c < \delta$, entonces,

$$V_f[a, b] - \tfrac{1}{2}\varepsilon \le \tfrac{1}{2}\varepsilon + \sum_{k=2}^{n} |f(x_k) - f(x_{k-1})| \le \tfrac{1}{2}\varepsilon + V_f[x_1, b],$$

por lo que se deduce que

$$V_f[c, x_1] = V_f[c, b] - V_f[x_1, b] < \varepsilon.$$

Demostrar que f es continua en $c \in [a, b]$ si y sólo si v_f es continua en c.

(f) Deducir que una función continua $f:[a, b] \to \mathbf{R}$ pertenece a $BV[a, b]$ si y sólo si es la diferencia de dos funciones continuas crecientes.

29.δ. Lebesgue demostró que una función con variación acotada, tiene derivada en todo punto, excepto posiblemente para un conjunto de "medida cero". La demostración de este resultado es bastante difícil y no se analizará aquí, pero se obtendrán otras propiedades de dichas funciones.

(a) Si $f \in BV[a, b]$ y si $c \in (a, b)$, entonces existen los límites por la derecha y por la izquierda de f en c. Estos límites son iguales, excepto, posiblemente para una colección numerable de puntos en $[a, b]$.

(b) Si (f_n) es una sucesión de funciones continuas en $BV[a, b]$ uniformemente convergente en $[a, b]$ a una función f, probar que no se deduce que f pertenezca a $BV[a, b]$.

(c) Sea (f_n) una sucesión en $BV[a, b]$ que converge en todo punto de $[a, b]$ a una función f y supóngase que para alguna $M > 0$se tiene $V_{f_n}[a, b] \le M$ para toda $n \in \mathbf{N}$. Demostrar que f pertenece a $BV[a, b]$ y que $V_f[a, b] \le M$.

(d) Sea (f_n) una sucesión en $BV[a, b]$ tal que $\|f_n - f_m\|_{BV} \to 0$ conforme $m, n \to \infty$. Demostrar que existe una función $f \in BV[a, b]$ tal que $\|f_n - f\|_{BV} \to 0$conforme $n \to \infty$.

(e) Sea (f_n)una sucesión de funciones monótonamente crecientes definidas en $I = [a, b]$ tal que $\|f_n\|_I \le M$ para toda $n \in \mathbf{N}$. Usar el método de la diagonal para obtener una subsucesión (g_k) de (f_n) que converja para cada número racional r en $[a, b]$. Defina $g(r) = \lim\,(g_k(r))$ para $r \in \mathbf{Q} \cap [a, b]$. Demostrar que g es creciente en $\mathbf{Q} \cap [a, b]$. Se define a g para $x \in [a, b)$ como el límite por la derecha $g(x) = \lim_{r \to x+} g(r)$. Demostrar que si $c \in [a, b)$ es un punto de continuidad, entonces $g(c) = \lim_k g_k(c)$. Dado que g tiene cuando más una colección numerable de puntos de discontinuidad, otra aplicación del método de la diagonal se puede aplicar para obtener una subsucesión (h_m) de (g_k) que converja en todo punto de $[a, b]$.

(f) Haciendo uso de la parte *(e)* demostrar el siguiente resultado, al que se llama **teorema de selección de Helly:** Sea (f_n) una sucesión de funciones en $BV[a, b]$ tal que $\|f_n\|_{BV} \leq M$ para toda $n \in N$. Entonces, existe una subsucesión de (f_n) que converge en todo punto de $[a, b]$ a una función $f \in BV[a, b]$ para la cual $\|f\|_{BV} \leq M$.

Sección 30 Existencia de la integral

En la sección anterior se demostraron algunas propiedades útiles de la integral de Riemann-Stieltjes. Sin embargo, aún no se ha demostrado la existencia de la integral para muchísimas funciones.

En esta sección se pondrá especial atención a funciones integrantes monótonamente crecientes, aunque gran parte de lo que aquí se haga se puede extender a funciones g que tengan variación acotada en un intervalo $J = [a, b]$ en el sentido de que exista una constante $M > 0$ tal que si $P = (x_0, x_1, \ldots, x_n)$ es cualquier partición de J entonces

$$(30.1) \qquad \sum_{j=1}^{n} |g(x_j) - g(x_{j-1})| \leq M.$$

Es claro que si g es monótonamente creciente, entonces, la suma en (30.1) se reduce y se puede tomar $M = g(b) - g(a)$. Por lo tanto, una función monótonamente creciente tiene variación acotada. Por el contrario, se puede demostrar que toda función con variación acotada es la diferencia de dos funciones crecientes. (Véase el proyecto 29.γ.)

Primero se demostrará un resultado muy eficaz.

30.1 CRITERIO DE RIEMANN PARA INTEGRABILIDAD. *Sea $J = [a, b]$ y sea g monótonamente creciente en J. Una función $f: J \to \mathbf{R}$ es integrable con respecto a g en J si y sólo si para toda $\varepsilon > 0$ existe una partición P_ε de J tal que si $P = (x_0, x_1, \ldots, x_n)$ es un refinamiento de P_ε, entonces,*

$$(30.2) \qquad \sum_{j=1}^{n} (M_j - m_j)\{g(x_j) - g(x_{j-1})\} < \varepsilon,$$

en donde $M_j = \sup\{f(x) : x \in [x_{j-1}, x_j]\}$ y $m_j = \inf\{f(x) : x \in [x_{j-1}, x_j]\}$ para $j = 1, \ldots, n$.

DEMOSTRACION. Si f es integrable con respecto a g y $\varepsilon > 0$ está dada, sea P_ε una partición de J tal que si $P = (x_0, x_1, \ldots, x_n)$ es un refinamiento de P_ε, entonces,

$$\left| S(P; f, g) - \int_a^b f\,dg \right| < \varepsilon$$

Para cualquier suma de Riemann-Stieltjes correspondiente a P. Ahora, elíjase y_j y z_j en $[x_{j-1}, x_j]$ tales que

$$M_j - \varepsilon < f(y_j), \qquad f(z_j) < m_j + \varepsilon.$$

Esto implica que $M_j - m_j < f(y_j) - f(z_j) + 2\varepsilon$ y por lo tanto

$$\sum_{j=1}^{n} (M_j - m_j)\{g(x_j) - g(x_{j-1})\} \leq \sum_{j=1}^{n} f(y_j)\{g(x_j) - g(x_{j-1})\}$$
$$- \sum_{j=1}^{n} f(z_j)\{g(x_j) - g(x_{j-1})\} + 2\varepsilon\{g(b) - g(a)\}.$$

El lado derecho de esta desigualdad contiene dos sumas de Riemann-Stieltjes correspondientes a P que no pueden diferir en más de 2ε. Por lo tanto, la condición (30.2) se satisface con ε reemplazada por $2\varepsilon\{1 + g(b) - g(a)\}$.

Por el contrario, suponga que $\varepsilon > 0$ está dada y que P_ε es una partición tal que (30.2) es válida para cualquier partición $P = (x_0, x_1, \ldots, x_n)$ que refine a P_ε. Sea $Q = (y_0, y_1, \ldots, y_m)$ un refinamiento de P; se va a calcular la diferencia $S(P; f, g) - S(Q; f, g)$ de dos sumas correspondientes. Dado que todo punto de P pertenece a Q se pueden expresar ambas sumas de la siguiente manera

$$S(P; f, g) = \sum_{k=1}^{m} f(u_k)\{g(y_k) - g(y_{k-1})\},$$

$$S(Q; f, g) = \sum_{k=1}^{m} f(v_k)\{g(y_k) - g(y_{k-1})\}.$$

Sin embargo, para expresar $S(P; f, g)$ en términos de los puntos en Q, se debe permitir la repetición de los puntos u_k y no exigir que u_k pertenezca a $[y_{k-1}, y_k]$. Ya que, tanto u_k como v_k pertenecen a algún intervalo $[x_{j-1}, x_j]$ y en este caso $|f(u_k) - f(v_k)| \leq M_j - m_j$. Multiplicando Por $g(y_k) - g(y_{k-1}) \geq 0$ y sumando se obtiene

$$|S(P; f, g) - S(Q; f, g)| \leq \sum (M_j - m_j)\{g(x_j) - g(x_{j-1})\} < \varepsilon.$$

Finalmente, sean P y P' refinamientos arbitrarios de P y sea Q un refinamiento común para P y P'. Dado que el argumento anterior es aplicable tanto a P como a P', se deduce que cualesquiera sumas $S(P; f, g)$ y $S(P'; f, g)$ podrían diferir en cuando más 2ε. Por lo tanto, el criterio de Cauchy se aplica para dar la integrabilidad de f.

Q.E.D.

30.2 TEOREMA DE INTEGRABILIDAD. Si f es continua y g es monótonamente creciente en J, entonces f es integrable con respecto a g en J.

DEMOSTRACION. Dado que f es uniformemente continua en J, dada $\varepsilon > 0$, existe una $\delta(\varepsilon) > 0$ tal que si $x, y \in J$ y $|x - y| < \delta(\varepsilon)$, entonces, $|f(x) - f(y)| < \varepsilon$. Sea $P_\varepsilon = (z_0, z_1, \ldots, z_r)$ una partición tal que sup $\{z_k - z_{k-1}\} < \delta(\varepsilon)$. Si $P = (x_0, x_1, \ldots, x_n)$ es un refinamiento de P_ε, entonces también sup $\{x_j - x_{j-1}\} < \delta(\varepsilon)$ de donde $M_j - m_j < \varepsilon$, por lo que se sigue que

$$\sum_{j=1}^{n} (M_j - m_j)\{g(x_j) - g(x_{j-1})\} \leq \varepsilon(g(b) - g(a)).$$

Dado que $\varepsilon > 0$ es arbitraria, el criterio de Riemann es aplicable.

Q.E.D.

30.3 COROLARIO. *Si f es monótona y g es continua en J, entonces, f es integrable con respecto a g en J.*

DEMOSTRACION. Aplicar el teorema anterior y el teorema 29.7 a $\pm f$.

Q.E.D.

El criterio de Riemann nos permite demostrar que el valor absoluto y el producto de funciones integrables son integrables.

30.4 TEOREMA. *Sea g monótonamente creciente en $J = [a, b]$.*

(a) Si $f : J \to \mathbf{R}$ es integrable, entonces, $|f|$ es integrable con respecto a g en J.

(b) Si f_1 y f_2 son integrables, entonces, el producto $f_1 f_2$ es integrable con respecto a g en J.

DEMOSTRACION. Defina M_j y m_j como en el criterio de Riemann y observe que

$$M_j - m_j = \sup \{f(x) - f(y) : x, y \in [x_{j-1}, x_j]\}.$$

Para demostrar *(a)*, observe que $||f(x)| - |f(y)|| \leq |f(x) - f(y)|$, por lo que el criterio de Riemann implica que $|f|$ es integrable cuando f lo es.

También se puede ver que si $|f(x)| \leq K$ para $x \in J$, entonces $|(f(x))^2 - (f(y))^2| \leq 2K |f(x) - f(y)|$, por lo que el criterio de Riemann implica que f^2 es integrable cuando f lo es. Para probar que $f_1 f_2$ es integrable cuando f_1 y f_2 lo son, observe que

$$2 f_1 f_2 = (f_1 + f_2)^2 - f_1^2 - f_2^2.$$

Q.E.D.

30.5 LEMA. *Sea g monótonamente creciente en $J = [a, b]$ y suponga que f es integrable con respecto a g en J. Entonces,*

$$\left| \int_a^b f \, dg \right| \leq \int_a^b |f| \, dg \leq \|f\|_J \, (g(b) - g(a)). \tag{30.3}$$

Si $m \leq f(x) \leq M$ para toda $x \in J$, entonces,

$$m(g(b) - g(a)) \leq \int_a^b f \, dg \leq M(g(b) - g(a)). \tag{30.4}$$

DEMOSTRACION. Del teorema 30.4 se deduce que $|f|$ es integrable con respecto a g. Si $P = (x_0, x_1, \ldots, x_n)$ es una partición de J y (z_j) es conjunto de puntos intermedios, entonces, para $j = 1, 2, \ldots, n$, se tiene

$$-\|f\|_J \leq -|f(z_j)| \leq f(z_j) \leq |f(z_j)| \leq \|f\|_J.$$

Multiplicar por $g(x_j) - g(x_{j-1}) \geq 0$ y sumar para obtener la estimación

$$-\|f\|_J\,(g(b)-g(a)) \leq -S(P;|f|,g) \leq S(P;f,g) \leq S(P;|f|,g)$$
$$\leq \|f\|_J\,(g(b)-g(a)),$$

por lo tanto, se sigue que $|S(P;f,g)| \leq S(P;|f|,g) \leq \|f\|_J\,(g(b)-g(a)$ lo que implica la validez de (30.3). La demostración de (30.4) es análoga y se omitirá.

Q.E.D.

Evaluación de la integral

Los dos resultados siguientes son útiles por sí mismos pero también nos llevan al teorema fundamental que es la herramienta principal para calcular integrables de Riemann.

30.6 PRIMER TEOREMA DEL VALOR MEDIO. *Si g es creciente en $J=[a,b]$ y f es continua de J a **R**, entonces, existe un número c en J tal que*

$$\int_a^b f\,dg = f(c)\int_a^b dg = f(c)\{g(b)-g(a)\}. \tag{30.5}$$

DEMOSTRACION. Si $m=\inf\{f(x):x\in J\}$ y $M=\sup\{f(x):x\in J\}$ en el lema anterior se vio que

$$m\{g(b)-g(a)\} \leq \int_a^b f\,dg \leq M\{g(b)-g(a)\}.$$

Si $g(b)=g(a)$, entonces, la relación (30.5) es trivial; si $g(b)>g(a)$, entonces, del teorema del valor intermedio de Bolzano 22.4 se sigue que existe un número c en J tal que

$$f(c)=\left\{\int_a^b f\,dg\right\}\Big/\{g(b)-g(a)\}.$$

Q.E.D.

30.7 TEOREMA DE DIFERENCIACION. *Suponga que f es continua en J y que g es creciente en J y tiene una derivada en un punto c en J. Entonces, la función F definida para x en J por medio de*

$$F(x)=\int_a^x f\,dg,$$

tiene una derivada en c y $F'(c)=f(c)g'(c)$.

DEMOSTRACION. Si $h>0$ es tal que $c+h$ pertenece a J, entonces, del teorema 29.6 y del resultado anterior se sigue que

$$F(c+h)-F(c)=\int_a^{c+h} f\,dg-\int_a^c f\,dg=\int_c^{c+h} f\,dg$$

$$= f(c_1)\{g(c+h) - g(c)\},$$

para alguna c_1 con $c \leq c_1 \leq c+h$. Una relación análoga es válida si $h < 0$. Dado que f es continua y g tiene una derivada en c, entonces, $F'(c)$ existe y es igual a $f(c)g'(c)$. Q.E.D.

Aplicando este teorema al caso de Riemann, se obtiene el resultado que proporciona las bases para el método de la evaluación de integrales en cálculo.

30.8 TEOREMA FUNDAMENTAL DEL CALCULO INTEGRAL. *Sea f continua en $J = [a, b]$. Una función F en J satisface*

$$F(x) - F(a) = \int_a^x f \quad \text{para } x \in J, \tag{30.6}$$

si y sólo si $F' = f$ en J.

DEMOSTRACION. Si la relación (30.6) es válida y $c \in J$, entonces, por el teorema anterior se puede ver que $F'(c) = f(c)$.

Por el contrario, defina a F_a para x en J como

$$F_a(x) = \int_a^x f.$$

El teorema anterior asegura que $F_a' = f$ en J. Si F es tal que $F' = f$, entonces, del teorema 27.9 (ii), se sigue que existe una constante C tal que $F(x) = F_a(x) + C$, para $x \in J$. Dado que $F_a(a) = 0$, entonces, $C = F(a)$ por lo que se deduce que si $F' = f$ en J, entonces,

$$F(x) - F(a) = \int_a^x f. \qquad \text{Q.E.D.}$$

NOTA. Si F es una función definida en **J** tal que $F' = f$ en J, algunas veces se dice que F es una **integral indefinida,** una **anti-derivada** o una **primitiva** de f. Con esta terminología, el teorema de diferenciación 30.7 asegura que toda función continua tiene una primitiva. Algunas veces, el teorema fundamental del cálculo integral se formula de maneras distintas a la que se da en 30.8, pero siempre incluye la afirmación de que, bajo hipótesis convenientes, la integral de Riemann de f se puede calcular evaluando cualquier primitiva de f en los punto puntos terminales del intervalo de integración. Se ha dado la formulación anterior que ofrece una condición necesaria y suficiente para que una función sea una primitiva de una función continua. Un resultado un poco más general que no requiere de la continuidad del integrando se encontrará en el ejercicio 30.J.

No se debe suponer que el teorema fundamental asegura que si la derivada f de una función F existe en todo punto de J, entonces, f es integrable y (30.6) es válido. De hecho, puede suceder que f no sea Riemann integrable

(véase el ejercicio 30.K.) En forma análoga una función f puede ser Riemann integrable pero no tener una primitiva (véase el ejercicio 30.L.).

Como consecuencia del teorema fundamental y del teorema 29.8, se obtiene la siguiente variante del primer teorema del valor medio 30.6, establecida aquí para integrales de Riemann.

30.9 PRIMER TEOREMA DEL VALOR MEDIO. *Si f y p son continuas en $J = [a, b]$ y $p(x) \geq 0$ para toda $x \in J$, entonces, existe un punto $c \in J$ tal que*

$$\int_a^b f(x)p(x)\,dx = f(c)\int_a^b p(x)\,dx. \tag{30.7}$$

DEMOSTRACION Defina $g: J \to \mathbf{R}$ para $x \in J$ por medio de

$$g(x) = \int_a^x p(t)\,dt.$$

Dado que $p(x) \geq 0$, se puede ver que g es creciente y se deduce del teorema de diferenciación 30.7 que $g' = p$. Por el teorema 29.8, se concluye que

$$\int_a^b f\,dg = \int_a^b fp,$$

y del primer teorema del valor medio 30.6, se deduce que para alguna c en J

$$\int_a^b f\,dg = f(c)\int_a^b p. \qquad \text{Q.E.D.}$$

Como una segunda aplicación del teorema 29.8, se reformulará el teorema 29.7 que se refiere a integración por partes en una forma más tradicional. La demostración se le deja al lector.

30.10 INTEGRACION POR PARTES. Si f y g tienen derivadas continuas en $[a, b]$, entonces

$$\int_a^b fg' = f(b)g(b) - f(a)g(a) - \int_a^b f'g.$$

El siguiente resultado, con frecuencia es útil.

30.11 SEGUNDO TEOREMA DEL VALOR MEDIO. *(a) Si f es creciente y g es continua en $J = [a, b]$, entonces, existe un punto c en J tal que*

$$\int_a^b f\,dg = f(a)\int_a^c dg + f(b)\int_c^b dg. \tag{30.8}$$

(b) Si f es creciente y h es continua en J, entonces, existe un punto c en J tal que

$$\int_a^b fh = f(a)\int_a^c h + f(b)\int_c^b h. \tag{30.9}$$

(c) Si φ es no negativa y creciente y h es continua en J, entonces, existe un punto c en J tal que

$$\int_a^b \varphi h = \varphi(b)\int_c^b h.$$

DEMOSTRACION. Las hipótesis, junto con el teorema de integrabilidad 30.6 implican que g es integrable con respecto a f en J. Más aún, por el primer teorema del valor medio 30.2

$$\int_a^b g\,df = g(c)\{f(b)-f(a)\}.$$

Después de usar el teorema 29.7, que se refiere a integración por partes, se concluye que f es integrable con respecto a g y

$$\begin{aligned}\int_a^b f\,dg &= \{f(b)g(b)-f(a)g(a)\}-g(c)\{f(b)-f(a)\}\\ &= f(a)\{g(c)-g(a)\}+f(b)\{g(b)-g(c)\}\\ &= f(a)\int_a^c dg + f(b)\int_c^b dg,\end{aligned}$$

lo que demuestra la parte *(a)*. Para demostrar *(b)*, defina a g en J por medio de

$$g(x) = \int_a^x h,$$

de tal manera que $g' = h$. La conclusión se sigue entonces de la parte *(a)* usando el teorema 29.8. Para demostrar *(c)* defina a F igual a φ para x en $(a, b]$ y defina $F(a) = 0$. Se aplica ahora la parte *(b)* a F. Q.E.D.

A la parte *(c)* del teorema anterior frecuentemente se le llama la forma de Bonnet† del segundo teorema del valor medio. Es evidente que hay un resultado correspondiente para una función decreciente (cf. ejercicio 30.N.)

Cambio de variable

En seguida se demostrará un teorema que justifica la fórmula familiar relacionada con el "cambio de variable" en una integral de Riemann.

30.12 TEOREMA DEL CAMBIO DE VARIABLE. Sea la función definida de un intervalo $[\alpha, \beta]$ a $\boldsymbol{R}$ en una derivada continua y suponga que $a = \varphi(\alpha)$ y $b = \varphi(\beta)$. Si f es continua en el rango de φ, entonces,

†OSSIAN BONNET (1819-1892) se le conoce principalmente por su trabajo en geometría diferencial.

(30.10) $$\int_a^b f(x)\,dx = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt.$$

DEMOSTRACION. Sea $I = \varphi([\alpha, \beta])$ defina a F, como

$$F(\xi) = \int_a^\xi f(x)\,dx \quad \text{para } \xi \in I$$

y considere a la función H definida por $H(t) = F(\varphi(t))$ para $\alpha \leq t \leq \beta$. Observe que $H(\alpha) = F(a) = 0$. Si se diferencia con respecto a t y se usa el hecho de que $F' = f$ (¿por qué?), se obtiene

$$H'(t) = F'(\varphi(t))\varphi'(t) = f(\varphi(t))\varphi'(t).$$

Se aplica ahora el teorema fundamental para deducir que

$$\int_a^b f(x)\,dx = F(b) = H(\beta) = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\,dt. \qquad \text{Q.E.D.}$$

Modificación de la integral

El siguiente resultado con frecuencia sirve para reducir una integral de Riemann-Stieltjes a una integral de Riemann.

30.13 TEOREMA. Si g' existe y f y g' son Riemann integrables en $[a, b]$, entonces, f es Riemann-Stieltjes integrable con respecto a g y

(30.11) $$\int_a^b f\,dg = \int_a^b fg'.$$

DEMOSTRACION. Sea $M > 0$ tal que $|f(x)| \leq M$ para $x \in [a, b]$ y sea $\varepsilon > 0$. Del teorema 30.4 se sigue que fg' es Riemann integrable. Por lo tanto, existe una partición P_ε de $[a, b]$ tal que si $P = (x_0, x_1, \ldots, x_n)$ es algún refinamiento de P_ε y si $\xi_j \in [x_{j-1}, x_j]$ para $j = 1, \ldots, n$, entonces,

(30.12) $$\left| \sum_{j=1}^n f(\xi_j)g'(\xi_j)(x_j - x_{j-1}) - \int_a^b fg' \right| < \varepsilon.$$

Como g' es Riemann integrable, también se puede suponer (en virtud del criterio de Riemann 30.1) que P_ε se ha elegido de tal manera que

(30.13) $$\sum_{j=1}^n (M_j - m_j)(x_j - x_{j-1}) < \varepsilon,$$

en donde $M_j = \sup\{g'(x) : x \in [x_{j-1}, x_j]\}$ y $m_j = \inf\{g'(x) : x \in [x_{j-1}, x_j]\}$. Si se usa el teorema del valor medio 27.6, se obtienen puntos $\zeta_j \in (x_{j-1}, x_j)$ tales que

$$\begin{aligned}\left|\sum_{j=1}^{n} f(\xi_j)\{g(x_j)-g(x_{j-1})\}-\int_a^b fg'\right| \\ =\left|\sum_{j=1}^{n} f(\xi_j)g'(\zeta_j)(x_j-x_{j-1})-\int_a^b fg'\right| \\ \leq \left|\sum_{j=1}^{n} f(\xi_j)\{g'(\zeta_j)-g'(\xi_j)\}(x_j-x_{j-1})\right| \\ +\left|\sum_{j=1}^{n} f(\xi_j)g'(\xi_j)(x_j-x_{j-1})-\int_a^b fg'\right|.\end{aligned}$$

Ahora, como $|g'(\zeta_j)-g'(\xi_j)|\leq M_j-m_j$, se sigue de (30.12) y (3O.13) que la expresión anterior está dominada por

$$M\sum_{j=1}^{n}(M_j-m_j)(x_j-x_{j-1})+\varepsilon\leq(M+1)\varepsilon.$$

Dado que $\varepsilon>0$ y la elección de $\xi_j\in[x_{j-1}, x_j]$ son arbitrarias, se sigue que f es integrable con respecto a g y que (30.11) es válido.

Q.E.D.

OBSERVACION. La demostración se puede modificar para aplicarla al caso en que f es acotada y g es continua en $[a, b]$, y en el que g tiene una derivada excepto en un número finito de puntos en los que g' se puede definir de tal manera que g' y fg' sean Riemann integrables en $[a, b]$.

Ejercicios

30.A. Demostrar que una función acotada que tenga cuando más un número finito de discontinuidades es Riemann integrable.

3O.B. Si $f:[a, b]\to \boldsymbol{R}$ es discontinua en algún punto del intervalo, entonces, existe una función g monótonamente creciente tal que f no es g-integrable.

3O.C. Demostrar que el teorema de integrabilidad 30.2 es válido cuando g es una función de variación acotada en J.

3O.D. Dar un ejemplo de una función f que no sea Riemann integrable sobre J pero tal que $|f|$ y f^2 sean Riemann integrables sobre J.

30.E. Sea f positiva y continua en $J=[a, b]$ y sea $M=\sup\{f(x):x\in J\}$. Demostrar que

$$M=\lim_n\left(\int_a^b (f(x))^n\,dx\right)^{1/n}.$$

30.F. Demostrar que el primer teorema del valor medio 30.6 puede no ser válido si f no es continua.

30.G. Demostrar que el teorema de diferenciación 30.7 es válido si se supone que f es integrable en J con respecto a una función creciente g, que f es continua en c y que g es diferenciable en c.

30.H. Suponga que f es integrable con respecto a una función creciente g en $J=[a, b]$ y defina a F para $x\in J$ como

$$F(x)=\int_a^x f\,dg.$$

Demostrar que: *(a)* si g es continua en c, entonces, F es continua en c y *(b)* si f es positiva, entonces, F es creciente

30.I. Dar un ejemplo de una función Riemann integrable f en J tal que la función F definida para $x \in J$ por medio de

$$F(x) = \int_a^x f,$$

no tenga derivada en algunos puntos de J. ¿Puede usted encontrar una función integrable f tal que no sea continua en J?

30.J. Si f es Riemann integrable en $J = [a, b]$ y si $F' = f$ en J, entonces

$$F(b) - F(a) = \int_a^b f.$$

(Sugerencia: si $P = (x_0, x_1, \ldots, x_n)$ es una partición de J, escribir

$$F(b) - F(a) = \sum_{j=1}^{n} \{F(x_j) - F(x_{j-1})\}.$$

30.K. Defina a F como

$$F(x) = x^2 \operatorname{sen}(1/x^2), \qquad 0 < x \leq 1,$$
$$= 0, \qquad x = 0.$$

Entonces, F tiene una derivada en todo punto de I. Sin embargo, F' no es integrable en I y F no es la integral de su derivada.

30.L. Defina a f como $f(x) = [x]$ para $x \in [0, 2]$. Entonces, f es Riemann integrable en $[0, 2]$, pero no es la derivada de ninguna función.

30.M. En el primer teorema del valor medio 30.9, suponga que P es Riemann integrable (en lugar de continua). Demostrar que la conclusión sigue siendo válida.

30.N. Si φ es no negativa y decreciente y h es continua en $[a, b]$, entonces, existe un punto $\xi \in [a, b]$ tal que

$$\int_a^b \varphi h = \varphi(a) \int_a^{\xi} h.$$

30.O. Sea f continua en $I = [0, 1]$, sea $f_0 = f$, y defina a f_{n+1} como

$$f_{n+1}(x) = \int_0^x f_n(t)\, dt \qquad \text{para } n \in \mathbf{N}, x \in I.$$

Por inducción, demostrar que $|f_n(x)| \leq (M/n!)x^n \leq M/n!$, en donde $M = \sup\{|f(x)| : x \in I\}$. Se sigue que la sucesión (f_n) converge uniformemente en I a la función cero.

30.P. Si f es integrable con respecto a g en $J = [a, b]$, si φ es continua y estrictamente creciente en $[c, d]$, y si $\varphi(c) = a$, $\varphi(d) = b$, entonces, $f \circ \varphi$ es integrable con respecto a $g \circ \varphi$ y

$$\int_a^b f\, dg = \int_c^d (f \circ \varphi)\, d(g \circ \varphi).$$

30.Q. Si f es continua en $[a, b]$ y si

$$\int_a^b fh = 0$$

para todas las funciones continuas h, entonces, $f(x) = 0$ para toda x.

30.R. Si f es integrable en $[a, b]$ y si

$$\int_a^b fh = 0$$

para todas las funciones continuas h, entonces, $f(x) = 0$ para todos los puntos de continuidad de f.

30.S. Sea P continua y positiva en $[a, b]$ y sea $c > 0$. Si

$$p(x) \leq c\int_a^x p(t)\,dt$$

para toda $x \in [a, b]$, demostrar que $p(x) = 0$ para toda x.

30.T. Sea f continua y tal que $f(x) \geq 0$ para toda $x \in [a, b]$. Si g es estrictamente creciente en $[a, b]$ demostrar que

$$\int_a^b f\,dg = 0$$

si y sólo si $f(x) = 0$ para toda $x \in [a, b]$.

30.U. Demostrar que si g es estrictamente creciente en $[a, b]$, entonces, en el primer teorema del valor medio 30.6 se puede tomar $c \in (a, b)$. Hacer una modificación análoga de las partes *(a)* y *(b)* del segundo teorema del valor medio 30.11.

30.V. Evaluar las siguientes integrales de Riemann-Stieltjes. (Aquí, $x \mapsto [x]$ la función del entero máximo.)

(a) $\int_0^1 x\,d(x^3)$, (b) $\int_{-2}^2 x\,d(|x|)$,

(c) $\int_0^2 x^3\,d([x])$, (d) $\int_0^4 x^2\,d([x^2])$,

(e) $\int_0^\pi \cos x\,d(\operatorname{sen} x)$, (f) $\int_{-\pi}^\pi \cos x\,d(|\operatorname{sen} x|)$.

Proyectos

30. α. El propósito de este proyecto es desarrollar el logaritmo usando una integral y su definición. Sea $P = \{x \in \mathbf{R} : x > 0\}$.

(a) Si $x \in P$, defina $L(x)$ como

$$L(x) = \int_1^x \frac{1}{t}\,dt.$$

por lo que $L(1) = 0$. Demostrar que L es diferenciable y que $L'(x) = 1/x$.

(b) Demostrar que $L(x) < 0$ para $0 < x < 1$ y $L(x) > 0$ para $x > 1$. De hecho,

$$1 - 1/x < L(x) < x - 1 \quad \text{para } x > 0.$$

(c) Demostrar que $L(xy) = L(x) + L(y)$ para x, y en P. Por lo tanto, $L(1/x) = -L(x)$ para x en P. *(Sugerencia:* si $y \in P$, defina a L_1 en P como $L_1(x) = L(xy)$ y demostrar que $L_1' = L'$.)

(d) Demostrar que si $n \in \mathbf{N}$, entonces

$$\frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} < L(n) < 1 + \frac{1}{2} + \cdots + \frac{1}{n-1}.$$

(e) Demostrar que L es una función uno a uno que se aplica a P sobre todo $\mathbf{R}$. Si e denota al único número tal que $L(e) = 1$, usando el hecho de que $L'(1) = 1$, demostrar que $e = \lim((1 + 1/n)^n)$.

(f) Sea r cualquier número racional positivo, entonces, $\lim_{x \to +\infty} L(x)/x^r = 0$.

(g) Observe que

$$L(1+x) = \int_1^{1+x} \frac{dt}{t} = \int_0^x \frac{dt}{1+t}.$$

Escribir $(1+t)^{-1}$ como una serie geométrica finita para obtener

$$L(1+x) = \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k} x^k + R_n(x).$$

Demostrar que $|R_n(x)| \leq 1/(n+1)$ para $0 \leq x \leq 1$ y

$$|R_n(x)| \leq \frac{|x|^{n+1}}{(n+1)(1+x)} \quad \text{para} -1 < x < 0.$$

30. β. Este proyecto dessarrolla las funciones trigonométricas empezando con una integral.

(a) Defínase A para x en $\mathbf{R}$ como

$$A(x) = \int_0^x \frac{dt}{1+t^2}.$$

Entonces A es una función impar (es decir, $A(-x) = -A(x)$), es estrictamente creciente y está acotada por 2. Defina π como $\pi/2 = \sup\{A(x) : x \in \mathbf{R}\}$.

(b) Sea T el inverso de A, de tal manera que T es una función estrictamente creciente con dominio $(-\pi/2, \pi/2)$. Demostrar que T tiene una derivada y que $T' = 1 + T^2$.

(c) Defina C y S en $(-\pi/2, \pi/2)$ por medio de las fórmulas

$$C = \frac{1}{(1+T^2)^{1/2}}, \qquad S = \frac{T}{(1+T^2)^{1/2}}.$$

Por lo que C es par y S es impar en $(-\pi/2, \pi/2)$. Demostrar que $C(0) = 1$ y $S(0) = 0$ y $C(x) \to 0$ y $S(x) \to 1$ conforme $x \to \pi/2$.

(d) Demostrar que $C'(x) = -S(x)$ y $S'(x) = C(x)$ para x en $(-\pi/2, \pi/2)$. Por lo tanto, tanto C como S satisfacen la ecuación diferencial

$$h'' + h = 0$$

en el intervalo $(-\pi/2, \pi/2)$.

(e) Defina $C(\pi/2) = 0$ y $S(\pi/2) = 0$ y defina C, S, T fuera del intervalo $(-\pi/2, \pi/2)$ por medio de las ecuaciones

$$C(x+\pi) = -C(x), \qquad S(x+\pi) = -S(x),$$
$$T(x+\pi) = T(x).$$

Si esto se hace sucesivamente, entonces C y S están definidas para todo R y tienen período 2π. Análogamente, T está definida excepto en múltiplos impares de $\pi/2$ y tiene período π.

(f) Demostrar que las funciones C y S, como se definieron en R en la parte anterior, son diferenciables en todo punto de R y siguen satisfaciendo las relaciones

$$C' = -S, \qquad S' = C$$

en todos R.

30.γ. Este proyecto desarrolla la importante fórmula del producto de Wallis.† En dicha fórmula se tomará

$$S_n = \int_0^{\pi/2} (\operatorname{sen} x)^n \, dx.$$

(a) Si $n > 2$, entonces, $S_n = [(n-1)/n]S_{n-2}$. (*Sugerencia:* integrar por partes.

(b) Probar las fórmulas

$$S_{2n} = \frac{1 \cdot 3 \cdot 5 \cdots (2n-1)}{2 \cdot 4 \cdot 6 \cdots (2n)} \frac{\pi}{2}, \qquad S_{2n+1} = \frac{2 \cdot 4 \cdots (2n)}{1 \cdot 3 \cdot 5 \cdots (2n+1)}.$$

(c) Demostrar que la sucesión (S_n) es monótonamente decreciente. (*Sugerencia:* $0 \leq \operatorname{sen} x \leq 1$.)

(d) Defina W_n como

$$W_n = \frac{2 \cdot 2 \cdot 4 \cdot 4 \cdot 6 \cdot 6 \cdots (2n)(2n)}{1 \cdot 3 \cdot 3 \cdot 5 \cdot 5 \cdot 7 \cdots (2n-1)(2n+1)}.$$

Demostrar que $\lim(W_n) = \pi/2$. (Producto de Wallis.)

(e) Demostrar que $\lim\left((n!)^2 2^{2n}/(2n)!\sqrt{n}\right) = \sqrt{\pi}$.

30.δ. Este proyecto desarrolla la importante fórmula de Stirling‡ la cual estima la magnitud de n!.

(a) Comparando el área bajo la hipérbola $y = 1/x$ y el área de un trapezoide inscrito en ella, demostrar que

$$\frac{2}{2n+1} < \log\left(1 + \frac{1}{n}\right).$$

A partir de esto, demostrar que $e < (1 + 1/n)^{n+1/2}$.

(b) Demostrar que

†JOHN WALLIS (1616-1703), profesor de Savilia de geometría en Oxford durante sesenta años, fue un precursor de Newton. Ayudó a implantar las bases para el desarrollo del cálculo.

‡ JAMES STIRLING (1692-1770) fue un matemático inglés de la escuela Newtoniana. la fórmula que se le atribuye a Stirling en realidad fue demostrada anteriormente por ABRAHAM DE MOIVRE (1667-1754), un hugonote francés quien se estableció en Londres y fue amigo de Newton.

$$\int_1^n \log x\, dx = n \log n - n + 1 = \log (n/e)^n + 1.$$

Considere la figura F hecha de rectángulos con bases $[1, \frac{3}{2}]$, $[n - \frac{1}{2}, n]$ y alturas 2, $\log n$, respectivamente, y de trapezoides con bases $[k - \frac{1}{2}, k + \frac{1}{2}]$, $k = 2, 3, \ldots, n - 1$, y con alturas inclinadas que pasan por los puntos $(k, \log k)$. Demostrar que el área de F es

$$1 + \log 2 + \cdots + \log (n-1) + \tfrac{1}{2} \log n = 1 + \log (n!) - \log \sqrt{n}.$$

(c) Comparando las dos áreas de la parte *(b)*, demostrar que

$$u_n = \frac{(n/e)^n \sqrt{n}}{n!} < 1, \qquad n \in \mathbf{N}.$$

(d) Demostrar que la sucesión (u_n) es monótonamente creciente. *(Sugerencia:* considere u_{n+1}/u_n.)

(e) Considerando u_n^2/u_{2n}, y haciendo uso del resultado de la parte *(e)* del proyecto anterior, demostrar que $\lim (u_n) = (2\pi)^{-1/2}$.

(f) Obtener la fórmula de Stirling

$$\lim \left(\frac{(n/e)^n \sqrt{2\pi n}}{n!} \right) = 1.$$

Sección 31 Otras propiedades de la integral

En esta sección se estudiarán otras propiedades de la integral de Riemann-Stieltjes (y de la de Riemann) que con frecuencia son útiles.

Primero se considera la posibilidad de "tomar el límite bajo el signo de la integral", es decir, la integrabilidad del límite de una sucesión de funciones integrables.

Suponga que g es monótonamente creciente en un intervalo $J = [a, b]$ y que (f_n) es una sucesión de funciones que son integrables con respecto a g y que convergen en todo punto de J a una función f. Es bastante natural esperar que la función límite f sea integrable y que

$$\int_a^b f\, dg = \lim \int_a^b f_n\, dg. \tag{31.1}$$

Sin embargo, éste no es necesariamente el caso aun para funciones muy exactas.

EJEMPLO. Sea $J = [0, 1]$, sea $g(x) = x$, y defina f_n Para $n \geq 2$ como

$$\begin{aligned} f_n(x) &= n^2 x, & 0 \leq x \leq 1/n, \\ &= -n^2(x - 2/n), & 1/n \leq x \leq 2/n, \end{aligned}$$

$$= 0, \qquad\qquad 2/n \leq x \leq 1.$$

Es claro que para cada n la función f_n es continua en J y por lo tanto es integrable con respecto a g. (figura 31.1). Ya sea por medio de un cálculo directo o remitiéndose al significado de la integral como un área, se obtiene

$$\int_0^1 f_n(x)\,dx = 1, \qquad n \geq 2.$$

Además, la sucesión (f_n) converge en todo punto de J a 0; Por lo que la función límite f es idénticamente cero, es integrable y

$$\int_0^1 f(x)\,dx = 0.$$

Por lo tanto, la ecuación (31.1) no es válida en este caso aún cuando ambos lados tengan un significado.

Dado que la ecuación (31.1) es muy conveniente, nos preguntamos si existen algunas otras condiciones sencillas que la impliquen. En seguida se demuestra que si la convergencia es uniforme, entonces, esta relación es válida.

31.2 TEOREMA. *Sea g una función monótonamente creciente en J y sea (f_n) una sucesión de funciones integrables con respecto a g sobre J. Suponga que la sucesión (f_n) converge uniformemente en J a una función límite f. Entonces, f es integrable con respecto a g y*

$$\int_a^b f\,dg = \lim \int_a^b f_n\,dg. \tag{31.1}$$

DEMOSTRACION. Sea $\varepsilon > 0$ y sea N tal que $\|f_N - f\|_J < \varepsilon$. Sea P_N una partición de J tal que si P, Q son refinamientos de P_N, entonces,

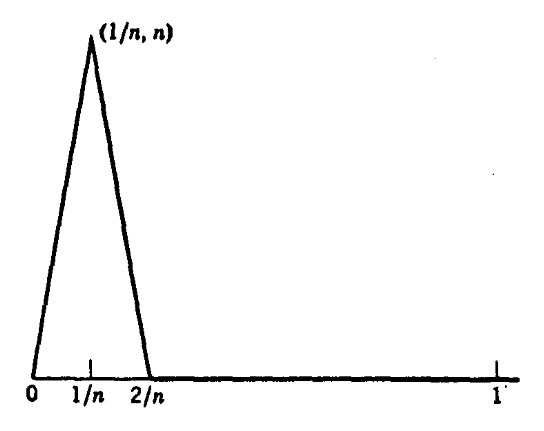


Figura 31.1 Gráfica de f_n.

$|S(P; f_N, g) - S(Q; f_N, g)| < \varepsilon$, para cualquier elección de los puntos intermedios. Si se usan los mismos puntos intermedios para f y f_N, entonces,

$$|S(P; f_N, g) - S(P; f, g)| \leq \sum_{k=1}^{n} \|f_N - f\|_J \{g(x_k) - g(x_{k-1})\}$$
$$= \|f_N - f\|_J \{g(b) - g(a)\} < \varepsilon\{g(b) - g(a)\}.$$

Dado que un cálculo similar es válido para la partición Q, entonces, para los refinamientos P, Q de P_N y sumas de Riemannn -Stieltjes correspondientes, se tiene

$$\begin{aligned}|S(P; f, g) - S(Q; f, g)| &\leq |S(P; f, g) - S(P; f_N, g)| \\ &\quad + |S(P; f_N, g) - S(Q; f_N, g)| \\ &\quad + |S(Q; f_N, g) - S(Q; f, g)| \\ &\leq \varepsilon(1 + 2\{g(b) - g(a)\}).\end{aligned}$$

De acuerdo con el criterio de Cauchy 29.4, la función límite f es integrable con respecto a g.

Para Probar (31.3) se usa el lema 30.5:

$$\left|\int_a^b f\,dg - \int_a^b f_n\,dg\right| = \left|\int_a^b (f - f_n)\,dg\right| \leq \|f - f_n\|_J \{g(b) - g(a)\}.$$

Y dado que $\lim \|f - f_n\|_J = 0$, se obtiene la conclusión deseada. Q.E.D.

La hipótesis supuesta en el teorema 31.2 de que la convergencia de la sucesión (f_n) es uniforme es un tanto severa y restringe la utilidad del resultado. Se establecerá ahora un resultado que no restringe tanto a la convergencia, pero que sí requiere de la integrabilidad de la función límite. Este resultado no se demostrará aquí, ya que la demostración más natural requeriría entrar a "teoría de la medida". (Sin embargo, el lector puede consultar el artículo de Luxemberg que aparece en las referencias.)

31.3 TEOREMA DE CONVERGENCIA ACOTADA. *Sea* (f_n) *una sucesión de funciones integrables con respecto a una función g monótonamente creciente de* $J = [a, b]$ *a* $\mathbf{R}$. *Suponga que existe* $B > 0$ *tal que* $\|f_n(x)\| \leq B$ *para toda* $n \in \mathbf{N}$, $x \in J$. *Si la función* $f(x) = \lim (f_n(x))$, $x \in J$, *existe y es integrable con respecto a g en J, entonces,*

$$\int_a^b f\,dg = \lim \int_a^b f_n\,dg. \tag{31.1}$$

La siguiente consecuencia del teorema de convergencia acotada con frecuencia es útil por lo que se demostrará formalmente.

31.4 TEOREMA DE CONVERGENCIA MONOTONA. *Sea* (f_n) *una sucesión monótona de funciones integrables con respecto a una función*

monótonamente creciente g *de* $J=[a,b]$ *a* R. *Si la función* $f(x)=\lim(f_n(x))$, $x\in J$, *es integrable con respecto a* g *en* J, *entonces,*

$$\int_a^b f\,dg=\lim\int_a^b f_n\,dg. \tag{31.1}$$

DEMOSTRACION. Suponga que $f_1(x)\leq f_2(x)\leq\cdots\leq f(x)$ Para toda $x\in J$. Entonces, $\|f_n(x)\|\leq B$, en donde $B=\|f_1\|_J+\|f\|_J$, Por lo que se puede aplicar 31.3.

Q.E.D.

El principal motivo de la importancia de la teoría de integración de Lebesgue (y Lebesgue-Stieltjes) es que amplía la clase de funciones integrables de tal manera que la ecuación (31.1) es válida bajo supuestos más débiles de los que se dan en teoremas anteriores. Ver *Los Elementos de Integración*, del autor, que aparece en las referencias.

Forma integral para el residuo

El lector recordará el teorema de Taylor 28.6' que permite calcular el valor $f(b)$ en términos de los valores $f(a), f'(a),\ldots,f^{(n-1)}(a)$ y un término residual que comprende a $f^{(n)}$ evaluada en un punto entre a y b. Para muchas aplicaciones es más conveniente poder expresar el término residual como una integral que comprenda a $f^{(n)}$.

31.5 TEOREMA DE TAYLOR. *Suponga que* f *y sus derivadas* $f', f'',\ldots,f^{(n)}$ *son continuas de* $[a,b]$ *a* R. *Entonces,*

$$f(b)=f(a)+\frac{f'(a)}{1!}(b-a)+\cdots+\frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1}+R_n,$$

en donde el residuo está dado por

$$R_n=\frac{1}{(n-1)!}\int_a^b(b-t)^{n-1}f^{(n)}(t)\,dt. \tag{31.2}$$

DEMOSTRACION. Integrar R_n por partes para obtener

$$R_n=\frac{1}{(n-1)!}\left\{(b-t)^{n-1}f^{(n-1)}(t)\Big|_{t=a}^{t=b}+(n-1)\int_a^b(b-t)^{n-2}f^{(n-1)}(t)\,dt\right\}$$
$$=-\frac{f^{(n-1)}(a)}{(n-1)!}(b-a)^{n-1}+\frac{1}{(n-2)!}\int_a^b(b-t)^{n-2}f^{(n-1)}(t)\,dt.$$

Si se continúa integrando por partes de esta manera, se obtiene la fórmula establecida.

Q.E.D.

En lugar de la fórmula (31.2), a menudo es conveniente hacer el cambio de variable $t=(1-s)a+sb$, Para s en $[0, 1]$, y obtener la fórmula

$$(31.3) \qquad R_n = \frac{(b-a)^{n-1}}{(n-1)!}\int_0^1 (1-s)^{n-1} f^{(n)}[a+(b-a)s]\, ds.$$

Esta forma del residuo se puede extender al caso en que f tiene dominio en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$.

Integrales dependientes de un parámetro

Con frecuencia es importante considerar integrales en las que los integrandos dependan de un parámetro. En tales casos, uno desea tener condiciones que aseguren la continuidad la diferenciabilidad y la integrabilidad de la función resultante. Los siguientes resultados son útiles en relación con esto.

Sea D el rectángulo en $\boldsymbol{R}\times\boldsymbol{R}$ dado Por

$$D=\{(x, t): a \le x \le b,\ c \le t \le d\},$$

y suponga que f es continua de D a $\boldsymbol{R}$. Entonces, se puede probar fácilmente (cf. ejercicio 22.G) que, para cada t fija en $[c, d]$, la función que manda a x a $f(x,t)$ es continua en $[a, b]$ y por lo tanto, Riemann integrable. Se define F para t en $[c, d]$ por medio de la fórmula

$$(31.4) \qquad F(t)=\int_a^b f(x, t)\, dx.$$

Primero se demostrará que F es continua.

31.6 TEOREMA. *Si f es continua de D a $\boldsymbol{R}$ y si F está definida por (31.4), entonces, F es continua de $[c, d]$ a $\boldsymbol{R}$.*

DEMOSTRACION. El teorema de continuidad uniforme 23.3 implica que si $\varepsilon>0$, entonces, existe una $\delta(\varepsilon)>0$ tal que si t y t_0 pertenecen a $[c, d]$ y $|t-t_0|<\delta(\varepsilon)$, entonces,

$$|f(x, t)-f(x, t_0)|<\varepsilon,$$

Para toda x en $[a, b]$. Del lema 30.5 se sigue que

$$\begin{aligned}|F(t)-F(t_0)| &= \left|\int_a^b \{f(x, t)-f(x, t_0)\}\, dx\right| \\ &\le \int_a^b |f(x, t)-f(x, t_0)|\, dx \le \varepsilon(b-a),\end{aligned}$$

lo que prueba la continuidad de F. Q.E.D

En los dos resultados siguientes, se usará el concepto de derivada parcial de una función de dos variables reales. Este concepto, familiar para el lector (por cálculo), se analizará más adelante, en el capítulo VII.

31.7 TEOREMA. *Si f y su derivada parcial fx son continuas de D a **R**, entonces, la función F definida por (31.4) tiene una derivada en* $[c, d]$ *y*

$$F'(t) = \int_a^b f_t(x, t)\, dx. \tag{31.5}$$

DEMOSTRACION. Por la continuidad uniforme de f_t en D se deduce,que $\varepsilon > 0$, entonces, hay una $\delta(\varepsilon) > 0$ tal que.si $|t - t_0| < \delta(\varepsilon)$, entonces,

$$|f_t(x, t) - f_t(x, t_0)| < \varepsilon$$

Para toda x en $[a, b]$. Sean t, t_0 tales que satisfagan esta condición, aplíquese el teorema del valor medio 27.6 para obtener una t_1 (que puede depender de x y esté entre t y t_0) tal que

$$f(x, t) - f(x, t_0) = (t - t_0) f_t(x, t_1).$$

Combinando estas dos relaciones, se deduce que si $0 < |t - t_0| < \delta(\varepsilon)$, entonces,

$$\left| \frac{f(x, t) - f(x, t_0)}{t - t_0} - f_t(x, t_0) \right| < \varepsilon,$$

para toda x en $[a, b]$. Aplicando el lema 30.5 se obtiene la estima

$$\left| \frac{F(t) - F(t_0)}{t - t_0} - \int_a^b f_t(x, t_0)\, dx \right| \le \int_a^b \left| \frac{f(x, t) - f(x, t_0)}{t - t_0} - f_t(x, t_0) \right| dx$$
$$\le \varepsilon(b - a),$$

que prueba la ecuación (31.5). Q.E.D.

Algunas veces, el parámetro t entra en los límites de integración así como en el integrando. En el siguiente resultado se considera esta posibilidad. En su demostración, se hará uso de un caso muy especial de la regla de la cadena (que se verá en el capítulo VII) la cual le será familiar al lector.

31.8 FORMULA DE LEIBNIZ. *Suponga que f y* f_t *son continuas de D a **R**, que* α *y* β *son funciones diferenciales en el intervalo* $[c, d]$ *y que tiene valores en* $[a, b]$. *Si* φ *está definida en* $[c, d]$ *por medio de*

$$\varphi(t) = \int_{\alpha(t)}^{\beta(t)} f(x, t)\, dx, \tag{31.6}$$

entonces φ *tiene una derivada para cada t en* $[c, d]$ *que está dada por*

(31.7) $$\varphi'(t) = f(\beta(t), t)\beta'(t) - f(\alpha(t), t)\alpha'(t) + \int_{\alpha(t)}^{\beta(t)} f_t(x, t)\, dx.$$

DEMOSTRACION. Defínase H para (u,v,t) como

$$H(u, v, t) = \int_v^u f(x, t)\, dx,$$

cuando u,v pertenecen a $[a, b]$ y t pertenece a $[c, d]$. La función φ definida en (31.6), es la composición dada por $\varphi(t) = H(\beta(t), \alpha(t), t)$. Aplicando la regla de la cadena se tiene

$$\varphi'(t) = H_u(\beta(t), \alpha(t), t)\beta'(t) + H_v(\beta(t), \alpha(t), t)\alpha'(t) + H_t(\beta(t), \alpha(t), t).$$

De acuerdo con el teorema de diferenciación 30.7,

$$H_u(u, v, t) = f(u, t), \qquad H_v(u, v, t) = -f(v, t),$$

y por el teorema anterior, se tiene

$$H_t(u, v, t) = \int_v^u f_t(x, t)\, dx.$$

Si se substituyen $u = \beta(t)$ y $v = \alpha(t)$, se obtiene la fórmula (31.7) Q.E.D.

Si f es continua de D a $\mathbf{R}$ y si F está definida por la fórmula (31.4), entonces se demostró que el teorema 31.6 que F es continua y por consecuencia Riemann integrable en el intervalo $[c, d]$. En seguida se demuestra que esta hipótesis de continuidad es suficiente para asegurar que se puede intercambiar el orden de integración. En fórmulas, esto se puede expresar como

(31.8) $$\int_c^d \left\{\int_a^b f(x, t)\, dx\right\} dt = \int_a^b \left\{\int_c^d f(x, t)\, dt\right\} dx.$$

31.9 TEOREMA DE INTERCAMBIO. *Si f es continua en D con valores en $\mathbf{R}$, entonces, la fórmula (31.8) es válida.*

DEMOSTRACION. El teorema 31.6 y el teorema de integrabilidad 30.2 implican que las dos integrales iteradas que aparecen en (31.8) existen; sólo queda demostrar la igualdad. Dado que f es uniformemente continua en D, si $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si $|x - x'| < \delta(\varepsilon)$ y $|t - t'| < \delta(\varepsilon)$, entonces, $|f(x, t) - f(x', t')| < \varepsilon$. Elija a n de tal magnitud que $(b - a)/n < \delta(\varepsilon)$ y $(d - c)/n < \delta(\varepsilon)$ y divida a D en rectángulos iguales dividiendo $[a, b]$ y $[c, d]$ en n partes iguales. Para $j = 0, 1, \ldots, n$, sea

$$x_j = a + (b - a)j/n, \qquad t_j = c + (d - c)j/n.$$

Se Puede escribir la integral de la izquierda en (31.8) como la suma

$$\sum_{k=1}^{n} \sum_{j=1}^{n} \int_{t_{k-1}}^{t_k} \left\{\int_{x_{j-1}}^{x_j} f(x, t)\, dx\right\} dt.$$

Aplicando el primer teorema del valor medio 30.6 dos veces, se deduce que existe un número x'_j en $[x_{j-1}, x_j]$ y un número t'_k en $[t_{k-1}, t_k]$ tal que

$$\int_{t_{k-1}}^{t_k} \left\{ \int_{x_{j-1}}^{x_j} f(x, t)\, dx \right\} dt = f(x'_j, t'_k)(x_j - x_{j-1})(t_k - t_{k-1}).$$

Por lo tanto, se tiene

$$\int_c^d \left\{ \int_a^b f(x, t)\, dx \right\} dt = \sum_{k=1}^{n} \sum_{j=1}^{n} f(x'_j, t'_k)(x_j - x_{j-1})(t_k - t_{k-1}).$$

La mísma línea de razonamiento aplicada a la integral de la derecha en (31.8), da la existencia de números x''_j en $[x_{j-1}, x_j]$ y t''_k en $[t_{k-1}, t_k]$ tales que

$$\int_a^b \left\{ \int_c^d f(x, t)\, dt \right\} dx = \sum_{k=1}^{n} \sum_{j=1}^{n} f(x''_j, t''_k)(x_j - x_{j-1})(t_k - t_{k-1}).$$

Dado que x'_j y x''_j pertenecen a $[x_{j-1}, x_j]$ y t'_k, t''_k pertenecen a $[t_{k-1}, t_k]$, Por la continuidad uniforme de f se concluye que las dos sumas dobles y por lo tanto las dos integrales iteradas, difieren, cuando más, en $\varepsilon(b-a)(d-c)$. Dado que $\varepsilon > 0$ es arbitraria, la igualdad de estas integrales se confirma.

Q.E.D.

Teorema de representación de Riesz

Esta sección concluirá con un teorema complejo que, a pesar de que no se aplicará más adelante, desempeña un papel muy importante en análisis funcional.

Primero, será conveniente enunciar algunos resultados que ya se han demostrado o que son consecuencia directa de lo que se ha hecho.

Sea $J = [a, b]$ una celda cerrada en R, denote por $C(J)$ el espacio vectorial de todas las funciones continuas de J a R y sea $C(J)$ la norma en $C(J)$ definida Por

$$\|f\|_J = \sup \{|f(x)| : x \in J\}.$$

Una **funciónal lineal** en $C(J)$ es una función lineal $G : C(J) \to R$ definida en el espacio vectorial $C(J)$; Por lo que

$$G(\alpha f_1 + \beta f_2) = \alpha G(f_1) + \beta G(f_2)$$

para todas α, β en R y f_1, f_2 en $C(J)$. Una funcional lineal G en $C(J)$ se dice que es **positiva** si para cada $f \in C(J)$ con $f(x) \geq 0$, $x \in J$, se tiene

$$G(f) \geq 0.$$

Una funcional lineal G en $C(J)$ se dice que es **acotada** si existe $M \geq 0$ tal que

$$|G(f)| \leq M \|f\|_J$$

† En una primera lectura el resto de esta sección se puede omitir.

para toda $f \in C(J)$.

31.10 LEMA. *Si g es una función monótonamente creciente en J y si G está definida para f en C(J) por medio de*

$$G(f) = \int_a^b f\,dg, \tag{31.9}$$

entonces G es una funcional lineal positiva acotada en C(J).

DEMOSTRACION. De los teoremas 29.5 *(a)* y 30.2 se deduce que G es una función lineal en $C(J)$ y del lema 30.5 que G es acotada con $M = g(b) - g(a)$. Si f pertenece a $C(J)$ y $f(x) \geq 0$ para $x \in J$, entonces, tomando $m = 0$ en la fórmula (30.4) se concluye que $G(f) \geq 0$.

Q.E.D.

Se demostrará ahora que, por el contrario, toda funcional lineal positiva acotada en $C(J)$ está generada por la integral de Riemann-Stieltjes con respecto a alguna función monótonamente creciente. Esta es una forma del célebre "teorema de representación de Riesz" que es una de las claves para el tema "análisis funcional" y tiene muchas generalizaciones y aplicaciones trascendentes. El teorema lo demostró el gran matemático húngaro Frederic Riesz.†

31.11 TEOREMA DE REPRESENTACION DE RIESZ. *Si G es una funcional lineal positiva acotada en C(J), entonces, existe una función g monótonamente creciente en J tal que*

$$G(f) = \int_a^b f\,dg, \tag{31.9}$$

para toda f en C(J).

DEMOSTRACION. Primero se definirá una función g monótonamente creciente y después se demostrará que (31.9) es válido.

Existe una constante M tal que si $0 \leq f_1(x) \leq f_2(x)$ para toda x en J, entonces, $0 \leq G(f_1) \leq G(f_2) \leq M\,\|f_2\|_J$. Si t es cualquier número real tal que $a < t < b$, y si n es un número natural lo suficientemente grande, se toma a como la función (figura 31.2) en $C(J)$ definida por

$$\begin{aligned} \varphi_{t,n}(x) &= 1, && a \leq x \leq t, \\ &= 1 - n(x - t), && t < x \leq t + 1/n, \\ &= 0, && t + 1/n < x \leq b. \end{aligned} \tag{31.10}$$

Es fácil ver que si $n \leq m$, entonces, para cada t con $a < t < b$,

† FREDERIC RIESZ (1880-1955), un brillante matemático húngaro, fué uno de los fundadores de topología y análisis funcional. También hizo valiosas aportaciones a la teoría potencial, ergódica y de integración.

$$0 \leq \varphi_{t,m}(x) \leq \varphi_{t,n}(x) \leq 1,$$

de tal manera que la sucesión $(G(\varphi_{t,n}) : n \in \mathbf{N})$ es una sucesión de números reales decreciente y acotada que converge a un número real. Se define $g(t)$ igual a este límite. Si $a < t \leq s < b$ y $n \in \mathbf{N}$, entonces,

$$0 \leq \varphi_{t,n}(x) \leq \varphi_{s,n}(x) \leq 1,$$

por lo que se deduce que $g(t) \leq g(s)$. Se define $g(a) = 0$ y si $\varphi_{b,n}$ denota la función $\varphi_{b,n}(x) = 1$, $x \in J$, entonces, se fija $g(b) = G(\varphi_{b,n})$. Si $a < t < b$ y n es lo suficientemente grande, entonces, para toda x en J se tiene

$$0 \leq \varphi_{t,n}(x) \leq \varphi_{b,n}(x) = 1,$$

y $g(a) = 0 \leq G(\varphi_{t,n}) \leq G(\varphi_{b,n}) = g(b)$. Esto prueba que $g(a) \leq g(t) \leq g(b)$ y completa la construcción de la función g monótonamente creciente.

Si f es continua en J y $\varepsilon > 0$, hay una $\delta(\varepsilon) > 0$ tal que si $|x - y| < \delta(\varepsilon)$ y $x, y \in J$, entonces, $|f(x) - f(y)| < \varepsilon$. Dado que f es integrable con respecto a g, existe una partición P_ε de J tal que si Q es un refinamiento de P_ε, entonces, para cualquier suma de Riemann-Stieltjes se tiene

$$\left| \int_a^b f\,dg - S(Q; f, g) \right| < \varepsilon.$$

Ahora, sea $P = (t_0, t_1, \ldots, t_m)$ una partición de J, en puntos distintos, que es un refinamiento de P_ε tal que sup $\{t_k - t_{k-1}\} < \frac{1}{2}\delta(\varepsilon)$ y sea n un número natural tan grande que

$$2/n < \inf\{t_k - t_{k-1}\}.$$

Entonces, sólo los intervalos consecutivos

(31.11) $$[t_0, t_1 + 1/n], \ldots, [t_{k-1}, t_k + 1/n], \ldots, [t_{m-1}, t_m]$$

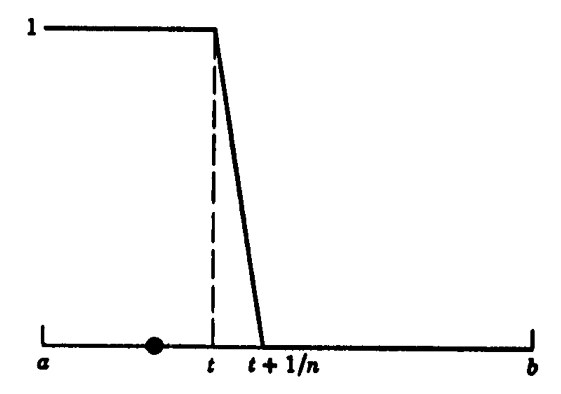


Figura 31.2 Gráfica de $\varphi_{t,n}$.

i279 (1478x2435x2 pcx)

tienen puntos en común (figura 31.3). Para cada $k = 1, \ldots, m$, la sucesión decreciente $(G(\varphi_{t_k,n}))$ converge a $g(t_k)$ y se puede suponer que n es tan grande que

(31.12) $$g(t_k) \le G(\varphi_{t_k,n}) \le g(t_k) + (\varepsilon/m \,\|f\|_J).$$

Considere ahora la función f^* definida en J como

(31.13) $$f^*(x) = f(t_1)\varphi_{t_1,n}(x) + \sum_{k=2}^{m} f(t_k)\{\varphi_{t_k,n}(x) - \varphi_{t_{k-1},n}(x)\}.$$

Un elemento x en J pertenece a un intervalo o bien a dos en (31.11). Si pertenece a un intervalo, entonces, se debe tener $t_0 \le x < t_1$ y $f^*(x) = f(t_1)$ o se tiene $t_{k-1} + (1/n) < x \le t_k$ para alguna $k = 1, 2, \ldots, m$ en cuyo caso $f^*(x) = f(t_k)$. (figura 31.4). Por lo tanto

$$|f(x) - f^*(x)| < \varepsilon.$$

Si x pertenece a dos intervalos en (31.11), entonces, $t_k \le x \le t_k + 1/n$ para alguna $k = 1, \ldots, m-1$ y se deduce que

$$f^*(x) = f(t_k)\varphi_{t_k,n}(x) + f(t_{k+1})\{1 - \varphi_{t_k,n}(x)\}.$$

Si se hace referencia a la definición de las φ's en (31.10), se tiene

$$f^*(x) = f(t_k)(1 - n(x - t_k)) + f(t_{k+1})n(x - t_k).$$

Dado que $|x - t_k| < \delta(\varepsilon)$ y $|x - t_{k+1}| < \delta(\varepsilon)$, se concluye que

$$\begin{aligned}|f(x) - f^*(x)| &\le |f(x) - f(t_k)|\,(1 - n(x - t_k)) + |f(x) - f(t_{k+1})|\,n(x - t_k)\\ &< \varepsilon\{1 - n(x - t_k) + n(x - t_k)\} = \varepsilon.\end{aligned}$$

Por consecuencia, se tiene el valor

$$\|f - f^*\|_J = \sup\{|f(x) - f^*(x)| : x \in J\} \le \varepsilon.$$

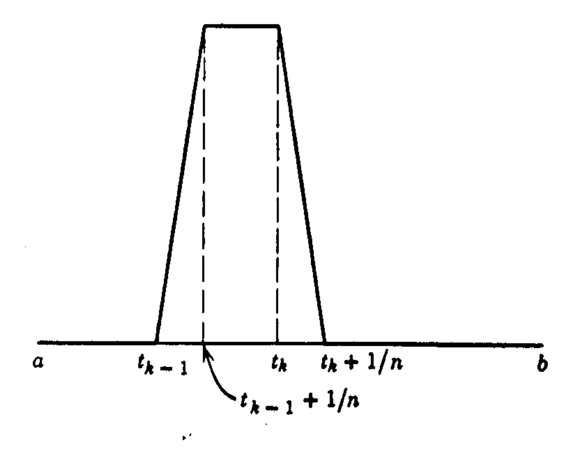


Figura 31.3 Gráfica de $\varphi_{t_k,n} - \varphi_{t_{k-1},n}$.

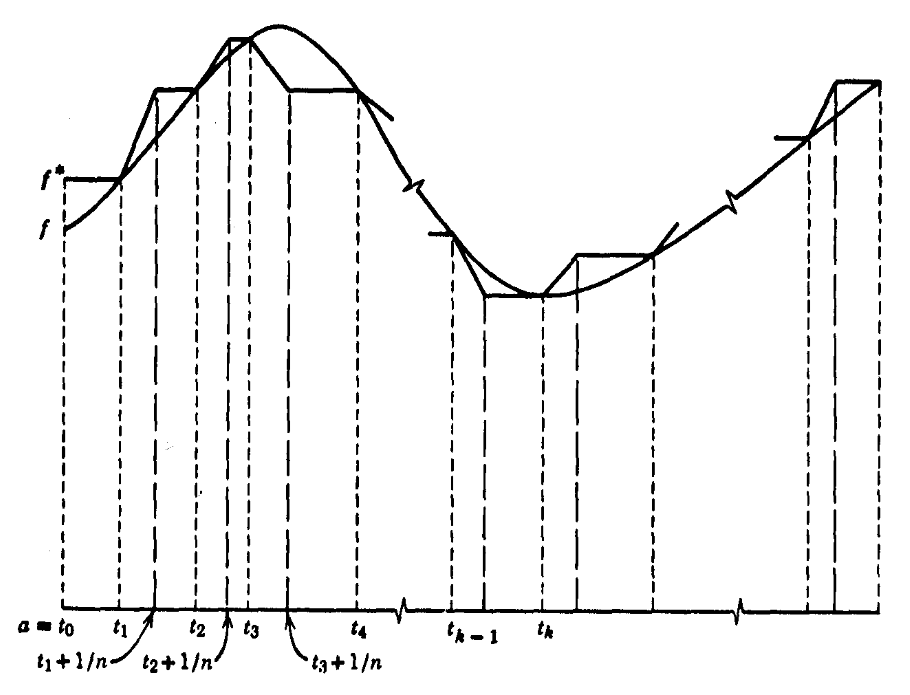


Figura 31.4 Gráficas de f y f_X

Dado que G es una función lineal en $C(J)$, se tiene que

$$|G(f) - G(f^*)| \leq M\varepsilon. \tag{31.14}$$

Por la relación (31.12) se ve que

$$|\{G(\varphi_{t_k,n}) - G(\varphi_{t_{k-1}n})\} - \{g(t_k) - g(t_{k-1})\}| < \varepsilon/2m\, \|f\|_J$$

para $k = 2, 3, \ldots, m$. Aplicando G a la función f^* definida por la ecuación (31.13) y recordando que $g(t_0) = 0$, se obtiene

$$\left| G(f^*) - \sum_{k=1}^{m} f(t_k)\{g(t_k) - g(t_{k-1})\} \right| < \varepsilon.$$

Pero el segundo término del lado izquierdo es una suma de Riemann-Stieltjes $S(P; f, g)$ para f con respecto a g correspondiente a la partición P que es un refinamiento de P_ε. Por lo tanto se tiene

$$\left| \int_a^b f\, dg - G(f^*) \right| \leq \left| \int_a^b f\, dg - S(P; f, g) \right| + |S(P; f, g) - G(f^*)| < 2\varepsilon.$$

Finalmente, usando la relación (31.14) se obtiene

$$\left| \int_a^b f\, dg - G(f) \right| < (M+2)\varepsilon. \tag{31.15}$$

Dado que $\varepsilon > 0$ es arbitraria y el lado izquierdo de (31.15) no depende de ella, se concluye que

$$G(f) = \int_a^b f\,dg. \qquad \text{Q.E.D.}$$

Para algunos propósitos, es importante saber que hay una correspondencia uno a uno entre funcionales lineales positivas acotadas en $C(J)$ y ciertas funciones normalizadas monótonamente crecientes. Esta construcción se puede checar para probar que da una función creciente g tal que $g(a) = 0$ y g es continua por la derecha en todo punto interior de J. Con estas condiciones adicionales, hay una correspondencia uno a uno entre funcionales positivas y funciones crecientes.

Ejercicios

31.A. Si $a > 0$, demostrar directamente que

$$\lim_n \int_0^a e^{-nx}\,dx = 0.$$

¿Cuáles de los resultados de esta sección son aplicables?

31.B. Si $0 < a < 2$, demostrar que

$$\lim_n \int_a^2 e^{-nx^2}\,dx = 0.$$

¿Qué sucede si $a = 0$?

31.C. Analizar $\lim_n \int_0^1 nx(1-x)^n\,dx$.

31.D. Si $a > 0$, demostrar que

$$\lim_n \int_a^\pi \frac{\operatorname{sen} nx}{nx}\,dx = 0.$$

¿Qué sucede si $a = 0$?

31.E. Sea $f_n(x) = nx(1+nx)^{-1}$ para $x \in [0, 1]$ y sea $f(x) = 0$ para $x = 0$ y $f(x) = 1$ para $x \in (0, 1]$. Demostrar que $f_n(x) \to f(x)$ para toda $x \in [0, 1]$ y que

$$\int_0^1 f_n(x)\,dx \to \int_0^1 f(x)\,dx.$$

31.F. Sea $h_n(x) = nx\,e^{-nx^2}$ para $x \in [0, 1]$ y sea $h(x) = 0$. Demostrar que

$$0 = \int_0^1 h(x)\,dx \neq \lim \int_0^1 h_n(x)\,dx = \frac{1}{2}.$$

31.G. Sea (g_n) una sucesión de funciones crecientes en $[a, b]$ que convergen uniformemente a una función g en $[a, b]$. Si una función creciente f es integrable con respecto a g_n para toda $n \in \mathbf{N}$, demostrar que f es integrable con respecto a g y

$$\int_a^b f\,dg = \lim \int_a^b f\,dg_n.$$

31.H. Dar un ejemplo para probar que la conclusión del ejercicio anterior puede no ser válida si la convergencia no es uniforme.

31.I. Si $\alpha > 0$, demostrar que $\int_0^1 t^\alpha (\log t)^2\, dt = 2/(\alpha+1)^3$.

31.J. Suponga que f y su derivada parcial f_t son continuas para (x,t) en $[a, b]\times[c, d]$. Aplicar el teorema de intercambio 31.9 a

$$\int_c^t \left\{\int_a^b f_t(x, t)\, dx\right\} dt, \qquad c \le t \le d,$$

y diferenciar para obtener otra demostración del teorema 31.7.

31.K. Usar el teorema fundamental 30.8 para demostrar que si una sucesión (f_n) de funciones converge en J a una función f y si las derivadas (f_n') son continuas y convergen uniformemente en J a una función g, entonces, f' existe y es igual a g. (Este resultado es menos general que el teorema 28.5 pero es más fácil de probar.)

31.L. Sea $\{r_1, r_2, \ldots, r_n, \ldots\}$ una enumeración de los números racionales en I. Defina a f_n como 1 si $x \in \{r_1, \ldots, r_n\}$ y como 0 en los otros casos. Entonces, f_n es Riemann integrable en I y la sucesión (f_n) converge monótonamente a la función discontinua f de Dirichlet (que es igual a 1 en $I \cap Q$ e igual a 0 en $I \setminus Q$.) por lo tanto, el límite monótono de una sucesión de funciones Riemann integrables no necesariamente es Riemann integrable.

31.M. Sea g una función fija monótonamente creciente en $J = [a, b]$. Si f es cualquier función integrable con respecto a g en J, entonces, se define $\|f\|_1$ como

$$\|f\|_1 = \int_a^b |f|\, dg.$$

Demostrar que las siguientes "propiedades de norma,, se cumplen:

(a) $\|f\|_1 \ge 0$;

(b) Si $f(x) = 0$ para toda $x \in J$, entonces, $\|f\|_1 = 0$;

(c) Si $c \in R$, entonces, $\|cf\|_1 = |c|\, \|f\|_1$;

(d) $|\, \|f\|_1 - \|h\|_1 | \le \|f \pm h\|_1 \le \|f\|_1 + \|h\|_1$. Sin embargo, es posible tener $\|f\|_1 = 0$ sin que se tenga $f(x) = 0$ para toda $x \in J$. (¿puede esto suceder cuando $g(x) = x$?)

31.N. Si g es monótonamente creciente en J y si f y f_n, $n \in N$, son funciones integrables con respecto a g, entonces, se dice que la sucesión (f_n) converge en la media (con respecto a g) siempre que

$$\|f_n - f\|_1 \to 0.$$

(La notación en este caso es la misma que en el ejercicio anterior.) Demostrar que si (f_n) converge en la media a f, entonces,

$$\int_a^b f_n\, dg \to \int_a^b f\, dg.$$

Demostrar que si una sucesion (f_n) de funciones integrables converge uniformemente en J a f, entonces, también converge en la media a f. De hecho,

$$\|f_n - f\|_1 \le \{g(b) - g(a)\}\, \|f_n - f\|_J.$$

Sin embargo, si f_n denota a la función del ejemplo 31.1 y si $g_n = (1/n) f_n$, entonces, la sucesión (g_n) converge en la media (con respecto a $g(x) = x$] a la función cero pero la convergencia no es uniforme en I.

31.O. Sea $g(x)=x$ en $J=[0, 2]$ y sea (I_n) una sucesión de intervalos cerrados en J tales que (i) la longitud de I_n es $1/n$, (ii) $I_n \cap I_{n+1}=\emptyset$, y (iii) todo punto x en J pertenece a una infinidad de los I_n. Defina f_n como

$$f_n(x)=1, \qquad x \in I_n,$$
$$=0, \qquad x \notin I_n.$$

Demostrar que la sucesión (f_n) converge en la media (con respecto a $g(x)=x$] a la función cero en J pero que la sucesión (f_n) no converge uniformemente. De hecho, la sucesión (f_n) no converge en *ningún* punto:

31.P. Sea g monótonamente creciente en $J=[a, b]$. Si f y h son integrables con respeco a g de J a $\boldsymbol{R}$, se define al **producto interno** (f,h) y f y h por medio de la fórmula

$$(f, h)=\int_a^b f(x)h(x)\, dg(x).$$

Verificar que todas las propiedades de la definición 8.3 se satisface excepto (ii). Si $f=h$ es la función cero en J, entonces, $(f, f)=0$; sin embargo, puede suceder que $(f, f)=0$ para una función f que no es cero en todo punto de J.

31.Q. Defina $\|f\|_2$ como

$$\|f\|_2=\left\{\int_a^b |f(x)|^2\, dg(x)\right\}^{1/2},$$

de manera que $\|f\|_2=(f, f)^{1/2}$. probar la desigualdad de Schwarz

$$|(f, h)| \leq \|f\|_2\, \|h\|_2$$

(teoremas 8.7 y 8.8). Probar que las propiedades de la norma 8.5 son válidas, sólo que $\|f\|_2=0$ no implica que $f(x)=0$ para toda x en J. Demostrar que $\|f\|_1 \leq \{g(b)-g(a)\}^{1/2}\, \|f\|_2$.

31.R. Sean f y f_n, $n \in \boldsymbol{N}$, integrables en J con respecto a una función creciente g. Se dice que la sucesión (f_n) converge en la media cuadrada (con respecto a g en J) a f si $\|f_n - f\|_2 \rightarrow 0$.

(a) Demostrar que si la sucesión es uniformemente convergente en, J entonces, también converge en la media cuadrada en la misma función.

(b) Demostrar que si la sucesión converge en media cuadrada, entonces, converge en la media a la misma función.

(c) Demostrar que el ejercicio 31.O prueba que la convergencia en la media cuadrada no implica la convergencia en cualquier punto de J.

(d) Si, en el ejercicio 31.O, se toma I_n de longitud $1/n^2$ y si se fija $h_n=nf_n$, entonces, la sucesión (h_n) converge en la media, pero no converge en la media cuadrada, a la función cero.

31.S. Probar que si la n-ésima derivada $f^{(n)}$ es continua en $[a, b]$, entonces, la forma integral del teorema de Taylor 31.5 y el primer teorema del valor medio 30.9 se pueden usar para obtener la forma de Lagrange del residuo dada en 28.6.

31.T. Si $J_1=[a, b]$, $J_2=[c, d]$, si f es continua de $J_1 \times J_2$ a $\boldsymbol{R}$ y g es Riemann integrable en J_1, entonces, la función F, definida en J_2 por medio de

$$F(t)=\int_a^b f(x, t)g(x)\, dx,$$

es continua en J_2.

31.U. Sea g una función creciente de $J_1=[a, b]$ a $\boldsymbol{R}$ y suponga que para cada t fija en $J_2=[c, d]$, la integral

$$F(t)=\int_a^b f(x, t)\, dg(x)$$

existe. Si la derivada parcial f_t es continua en $J_1 \times J_2$, entonces, la derivada F' existe en J_2 y está dada por

$$F'(t)=\int_a^b f_t(x, t)\, dg(x).$$

31.V. Sean $J_1=[a, b]$ y $J_2=[c, d]$. Suponga que la función de valor real g es monótona en J_1, que h es monótona en J_2, y que f es continua en $J_1 \times J_2$. Defina G en J_2 y H en J_1 como

$$G(t)=\int_a^b f(x, t)\, dg(x), \qquad H(x)=\int_c^d f(x, t)\, dh(t).$$

Demostrar que G es integrable con respecto a h en J_2, que H es integrable con respecto a g en J_1 y que

$$\int_c^d G(t)\, dh(t)=\int_a^b H(x)\, dg(x).$$

Esta última ecuación se puede escribir de la forma

$$\int_c^d\left\{\int_a^b f(x, t)\, dg(x)\right\} dh(t)=\int_a^b\left\{\int_c^d f(x, t)\, dh(t)\right\} dg(x).$$

31.W. Sean, f, J_1, y J_2 como en el ejercicio 31.V. Si φ está en $C(J_1)$ (es decir, φ es una función continua de J_1 a $\boldsymbol{R}$), sea $T(\varphi)$ la función definida en J_2 por medio de la fórmula

$$T(\varphi)(t)=\int_a^b f(x, t) \varphi(x)\, dx.$$

Demostrar que T es una **transformación lineal** de $C(J_1)$ a $C(J_2)$ en el sentido de que si φ, ψ pertenecen a $C(J_1)$, entonces,

(a) $T(\varphi)$ pertenece a $C(J_2)$,

(b) $T(\varphi+\psi)=T(\varphi)+T(\psi)$,

(c) $T(c\varphi)=cT(\varphi)$ para $c \in \boldsymbol{R}$. Si $M=\sup \{|f(x, t)|:(x, t) \in J_1 \times J_2\}$, entonces, T es **acotada** en el sentido de que

(d) $\|T(\varphi)\|_{J_2} \leq M\|\varphi\|_{J_1}$ para $\varphi \in C(J_1)$.

31.X. Continuando con la notación del ejercicio anterior, demostrar que si $r>0$, entonces, T manda a la colección

$$B_r=\{\varphi \in C(J_1):\|\varphi\|_{J_1} \leq r\}$$

hacia un conjunto uniformemente equicontinuo de funciones en $C(J_2)$ (véase la definición 28.6). Por lo tanto, si (φ_n)es cualquier sucesión de funciones en B_r, hay una subsucesión (φ_{n_k}) tal que la sucesión $(T(\varphi_{n_k}))$ converge uniformemente en J_2.

31.Y. Defínanse J_1 y J_2 como ya se hizo y sea f continua de $\boldsymbol{R}\times J_2$ a $\boldsymbol{R}$. Si φ está en $C(J_1)$, sea $S(\varphi)$ la función definida en J_2 por medio de la fórmula

$$S(\varphi)(t)=\int_a^b f(\varphi(x),t)\,dx.$$

Demostrar que $S(\varphi)$ pertenece a $C(J_2)$, pero que, en general, S no es una transformación lineal en el sentido del ejercicio 31.W. Sin embargo, demostrar que S manda a la colección B_r del ejercicio 31.X hacia el conjunto uniformemente equicontinuo de funciones en $C(J_2)$. Además, si (φ_n) es cualquier sucesión en B_r, existe una subsucesión tal que $(S(\varphi_{n_k}))$ converge uniformemente en J_2. (Este resultado es importante en la teoría de ecuaciones integrales no lineales.)

31.Z. Demostrar que si se definen G_0, G_1, G_2 para f en $C(\boldsymbol{I})$ como

$$G_0(f)=f(0),\qquad G_1(f)=2\int_0^{1/2}f(x)\,dx,$$
$$G_2(f)=\tfrac{1}{2}\{f(0)+f(1)\};$$

entonces, G_0, G_1, y G_2 son funcionales lineales positivas acotadas en $C(\boldsymbol{I})$. Dar funciones monótonamente crecientes g_0, g_1, g_2 que representen a estas funcionales lineales como integrales de Riemann-Stieltjes.

Demostrar que la elección de estas g_j no está determinada de manera única a menos de que se pida que $g_j(0)=0$ y que g_j sea continua por la derecha en cada punto interior de $\boldsymbol{I}$.

Proyecto

31.α. Este proyecto prueba la existencia de una solución única de una ecuación diferencial de primer orden bajo una condición de Lipschitz. Sea $\Omega\subseteq\boldsymbol{R}^2$ abierto y sea $f:\Omega\to\boldsymbol{R}$ continua tal que satisfaga la condición de Lipschitz: $|f(x,y)-f(x,y')|\le K\,|y-y'|$ para todos los puntos $(x,y),(x,y')$ en Ω. Sea $\boldsymbol{I}$ una celda cerrada

$$\boldsymbol{I}=\{(x,y):|x-a|\le\alpha,\ |y-b|\le\beta\}$$

contenida en Ω y suponga que $M\alpha\le\beta$, en donde $|f(x,y)|\le M$ para $(x,y)\in\boldsymbol{I}$.

(a) Si $J=[a-\alpha,a+\alpha]$, defina $\varphi_0(x)=b$ para $x\in J$ y, si $n\in\boldsymbol{N}$, defina

$$\varphi_n(x)=b+\int_a^x f(t,\varphi_{n-1}(t))\,dt$$

para $x\in J$. demostrar por inducción que la sucesión (φ_n) está bien definida en J y que

(i) $|\varphi_n(x)-b|\le\beta$,

(ii) $|\varphi_n(x)-\varphi_{n-1}(x)|\le\dfrac{MK^{n-1}}{(n-1)!}|x-a|^{n-1}$,

para toda $x\in J$.

(b) Demostrar que cada una de las funciones φ_n es continua en J y que la sucesión (φ_n) converge uniformemente en J a una función φ.

(c) Concluir que la función φ es continua en J, satisface $\varphi(a)=b$, y

$$\varphi(x) = b + \int_a^x f(t, \varphi(t))\, dt,$$

para toda $x \in J$. deducir que φ es diferenciable en J y satisface

$$\varphi'(x) = f(x, \varphi(x)) \qquad \text{para } x \in J.$$

(d) Si ψ es continua en J y satisface

$$\psi(a) = b, \qquad \psi'(t) = f(x, \psi(x))$$

para toda $x \in J$, demostrar que

$$\psi(x) = b + \int_a^b f(t, \psi(t))\, dt \qquad \text{para } x \in J.$$

(e) Si φ es como en *(c)* y ψ es como en *(d)*, demostrar por inducción que

$$|\varphi(x) - \psi(x)| \leq K \left| \int_a^x |\varphi(t) - \psi(t)|\, dt \right|$$
$$\leq \frac{K^n}{n!} \|\varphi - \psi\|_J \, |x - a|^n.$$

Por lo tanto $\|\varphi - \psi\|_J \leq \|\varphi - \psi\|_J K^n \alpha^n / n!$, por lo que se deduce que $\varphi(x) = \psi(x)$ para toda $x \in J$.

Sección 32 Integrales impropias e infinitas

En las tres secciones anteriores ha habido dos suposiciones básicas: se requiere que las funciones sean acotadas y se requiere que el dominio de integración sea compacto. Si alguna de estas hipótesis se elimina, la teoría de integración subsecuente no es aplicable sin algún cambio. Dado que hay varias aplicaciones importantes en donde es deseable que se permita una o las dos nuevas situaciones, en seguida se indicarán los cambios que se deben hacer.

Funciones no acotadas

Sea $J = [a, b]$ un intervalo en $\boldsymbol{R}$ y sea f una función de valor real definida al menos para las x que satisfacen $a < x \leq b$. Si f es Riemann integrable en el intervalo $[c, b]$ para cada c que satisfaga $a < c \leq b$, sea

(32.1) $$I_c = \int_c^b f.$$

Se definirá la integral impropia de f sobre $J = [a, b]$ como el límite de I_c cuando $c \to a$.

32.1 DEFINICION. Suponga que la integral de Riemann en (32.1) existe para cada c en $(a, b]$. Suponga que existe un número real I tal que para

toda $\varepsilon > 0$ existe una $\delta(\varepsilon) > 0$ tal que si $a < c < a + \delta(\varepsilon)$ entonces, $|I_c - I| < \varepsilon$. En este caso se dice que I es la **integral impropia** de f sobre $J = [a, b]$ y algunas veces se denota el valor I de esta integral impropia por medio de

$$\text{(32.2)} \qquad \int_{a+}^{b} f \qquad \text{o por} \qquad \int_{a+}^{b} f(x)\, dx,$$

aunque es más usual no escribir el signo de suma en el límite inferior.

32.2 EJEMPLOS. *(a)* Suponga que la función f está definida en $(a, b]$ y está acotada en este intervalo. Si f es Riemann integrable en todo intervalo $[c, b]$ con $a < c \leq b$, entonces, fácilmente se ve (ejercicio 32.A) que la integral impropia (32.2) existe. De modo que la función $f(x) = \text{sen}\,(1/x)$ tiene una integral impropia en el intervalo $[0, 1]$.

(b) Si $f(x) = 1/x$ para x en $(0, 1]$ y si c está en $(0, 1]$ entonces, del teorema fundamental 30.8 y del hecho de que f es la derivada del logaritmo se sigue que

$$I_c = \int_c^1 f = \log 1 - \log c = -\log c.$$

Dado que $\log c$ se hace no acotado conforme $c \to 0$, la integral impropia de f en $[0, 1]$ no existe.

(c) Sea $f(x) = x^\alpha$ para x en $(0, 1]$. Si $\alpha < 0$, la función es continua pero no acotada en $(0, 1]$. Si $\alpha \neq -1$, entonces, f es la derivada de

$$g(x) = \frac{1}{\alpha + 1} x^{\alpha + 1}.$$

Del teorema fundamental 30.8 se tiene que

$$\int_c^1 x^\alpha\, dx = \frac{1}{\alpha + 1}(1 - c^{\alpha + 1}).$$

Si α satisface $-1 < \alpha < 0$, entonces, $c^{\alpha+1} \to 0$ cuando $c \to 0$ y f tiene una integral impropia. Por otro lado, $\alpha < -1$, entonces, $c^{\alpha+1}$ no tiene un límite (finito) cuando $c \to 0$, y por lo tanto f no tiene una integral impropia.

El análisis anterior pertenece a una función que no está definida o no está acotada en el punto extremo izquierdo del intervalo. Resulta obvio cómo debe tratarse el caso análogo en el punto extremo derecho. El caso en el que la función no está definida o no está acotada en un punto interior del intervalo es un poco más interesante. Suponga que p es un punto interior de $[a, b]$ y que f está definida en todo punto de $[a, b]$ excepto, posiblemente p. Si las dos integrales impropias

$$\int_a^{p-} f, \qquad \int_{p+}^{b} f$$

existen, entonces, se define la **integral impropia** de f sobre $[a, b]$ como su suma. En la notación del límite, se define la integral impropia de f sobre $[a, b]$ como

(32.3) $$\lim_{\varepsilon\to 0+} \int_a^{p-\varepsilon} f(x)\, dx + \lim_{\delta\to 0+} \int_{p+\delta}^{b} f(x)\, dx.$$

Es claro que si esos dos límites existen, entonces, el límite

(32.4) $$\lim_{\varepsilon\to 0+} \left\{ \int_a^{p-\varepsilon} f(x)\, dx + \int_{p+\varepsilon}^{b} f(x)\, dx \right\}$$

también existe y tiene el mismo valor. Sin embargo, la existencia del límite (32.4) no implica la existencia de (32.3). Por ejemplo, si f está definida para $x \in [-1, 1]$, $x \neq 0$, por medio de $f(x) = 1/x^3$, entonces, es fácil ver que

$$\int_{-1}^{-\varepsilon} \left(\frac{1}{x^3}\right) dx + \int_{\varepsilon}^{1} \left(\frac{1}{x^3}\right) dx = \left(\frac{-1}{2}\right)\left(\frac{1}{\varepsilon^2} - 1\right) + \left(\frac{-1}{2}\right)\left(1 - \frac{1}{\varepsilon^2}\right) = 0$$

para toda ε que satisfaga $0 < \varepsilon < 1$. Sin embargo, en el ejemplo 32.2 *(c)* se vio que si $\alpha = -3$, entonces, las integrales impropias.

$$\int_{-1}^{0-} \frac{1}{x^3}\, dx, \qquad \int_{0+}^{1} \frac{1}{x^3}\, dx$$

no existen.

Los comentarios anteriores prueban que el límite en (32.4) puede existir sin que el límite en (32.3) exista. Se define la integral impropia (a la que algunas veces se le llama **integral de Cauchy**) de f como aquella dada por (32.3). El límite en (32.4) también es de interés y se le llama el **valor principal de Cauchy** de la integral y se denota por medio de

$$\text{(CPV)} \int_a^b f(x)\, dx.$$

Es claro que una función que tiene un número finito de puntos en donde no está definida o acotada se puede tratar partiendo el intervalo en subintervalos en donde estos puntos sean puntos extremos

Integrales infinitas

Es importante extender la integral a ciertas funciones que estén definidas en conjuntos no acotados. Por ejemplo, si f está definida de $\{x \in \boldsymbol{R} : x \geq a\}$ a $\boldsymbol{R}$ y es Riemann integrables sobre $[a, c]$ para toda $c > a$, se toma I_c como la integral parcial dada por

(32.5) $$I_c = \int_a^c f.$$

En seguida se definirá la "integral infinita" de f para $x \geq a$ como el límite de I_c cuando c aumenta.

32.3 DEFINICION Si f es Riemann integrable sobre $[a, c]$ para cada $c > a$, sea I_c la **integral parcial** dada por (32.5). Se dice que un número real I es la **integral infinita** de f sobre $\{x : x \geq a\}$ si para toda $\varepsilon > 0$, existe un número real $M(\varepsilon)$ tal que si $c > M(\varepsilon)$ entonces $|I - I_c| < \varepsilon$. En este caso se denota a I por medio de

(32.6) $$\int_a^{+\infty} f \quad \text{ó} \quad \int_a^{+\infty} f(x)\, dx.$$

Es necesario aclarar que a las integrales infinitas algunas veces se les llama "integrales impropias del primer tipo." Se ha elegido la terminología actual, debidas a Hardy† ya que es más simple y es paralela a la terminología que se usa en relación con series infinitas.

32.4 EJEMPLOS. *(a)* Si $f(x) = 1/x$ para $x > a > 0$, entonces, las integrales parciales son

$$I_c = \int_a^c \frac{1}{x}\, dx = \log c - \log a.$$

Dado que $\log c$ se hace no acotado cuando $c \to +\infty$, la integral infinita de f no existe.

(b) Sea $f(x) = x^\alpha$ para $x \geq a > 0$ y $\alpha \neq -1$. Entonces

$$I_c = \int_a^c x^\alpha\, dx = \frac{1}{\alpha + 1}(c^{\alpha+1} - a^{\alpha+1}).$$

Si $\alpha > -1$, entonces, $\alpha + 1 > 0$ y la integral infinita no existe. Sin embargo, si $\alpha < -1$, entonces,

$$\int_a^{+\infty} x^\alpha\, dx = -\frac{a^{\alpha+1}}{\alpha + 1}.$$

(c) Sea $f(x) = e^{-x}$ para $x \geq 0$. Entonces,

$$\int_0^c e^{-x}\, dx = -(e^{-c} - 1);$$

por lo que la integral infinita de f sobre $\{x : x \geq 0\}$ existe y es igual a 1.

También es posible considerar la integral de una función definida en todo $\boldsymbol{R}$. En este caso se requiere que f sea Riemann integrable sobre todo intervalo finito en $\boldsymbol{R}$ y se consideran los límites

(32.7a) $$\int_{-\infty}^a f(x)\, dx = \lim_{b \to -\infty} \int_b^a f(x)\, dx,$$

† GEOFFREY H. HARDY (1877-1947) fue profesor en Cambridge y durante mucho tiempo Decano de las matemáticas inglesas. Hizo muchas y profundas contribuciones al análisis matemático.

(32.7b) $$\int_a^{+\infty} f(x)\,dx = \lim_{c\to+\infty} \int_a^c f(x)\,dx.$$

Es fácil darse cuenta de que si ambos límites existen para un valor de a, entonces, existen para todos los valores de a. En este caso se define la **integral infinita de f sobre R** como la suma de estas dos integrales infinitas:

(32.8) $$\int_{-\infty}^{+\infty} f(x)\,dx = \lim_{b\to-\infty} \int_b^a f(x)\,dx + \lim_{c\to+\infty} \int_a^c f(x)\,dx.$$

Igual que en el caso de la integral impropia, la existencia de los dos límites en (32.8) implica la existencia del límite

(32.9) $$\lim_{c\to+\infty} \left\{ \int_{-c}^{a} f(x)\,dx + \int_a^c f(x)\,dx \right\} = \lim_{c\to+\infty} \int_{-c}^{c} f(x)\,dx,$$

y la igualdad de (32.8) y (32.9). Al límite en (32,9), cuando existe, a menudo se le llama el **valor principal de Cauchy** de la integral infinita sobre R y se denota por

(32.10) $$\text{(CPV)} \int_{-\infty}^{+\infty} f(x)\,dx.$$

Sin embargo, la existencia del valor principal de Cauchy no implica la existencia de la integral infinita (32.8). Esto se puede ver tomando $f(x) = x$, por lo que

$$\int_{-c}^{c} x\,dx = \tfrac{1}{2}(c^2 - c^2) = 0$$

para toda c. De modo que el valor principal de Cauchy de la integral infinita para $f(x) = x$ existe y es igual pero b la integral infinita de esta función no existe ya que ninguna de las integrales infinitas en (32.7) existe.

Existencia de la integral infinita

Se obtendrán ahora algunas condiciones para la existencia de la integral infinita sobre el conjunto $\{x : x \geq a\}$. Estos resultados también se pueden aplicar para dar condiciones para la integral infinita sobre R ya que esta última implica la consideración de integrales infinitas sobre los conjuntos $\{x : x \leq a\}$ y $\{x : x \geq a\}$. Primero se probará el criterio de Cauchy.

32.5 CRITERIO DE CAUCHY. *Suponga que f es integrable sobre $[a, c]$ para toda $c \geq a$. Entonces la integral infinita*

$$\int_a^{+\infty} f$$

existe si y sólo si para toda $\varepsilon > 0$ existe una $K(\varepsilon)$ tal que si $b \geq c \geq K(\varepsilon)$, entonces,

(32.11) $$\left| \int_c^b f \right| < \varepsilon.$$

DEMOSTRACION. La necesidad de la condición se prueba de la manera usual. Suponga que la condición se satisface y sea I_n la integral parcial definida para $n \in N$ como

$$I_n = \int_a^{a+n} f.$$

Se puede ver que (I_n) es una sucesión de Cauchy de números reales. Si $I = \lim (I_n)$ y $\varepsilon > 0$, entonces, existe $N(\varepsilon)$ tal que si $n \geq N(\varepsilon)$, entonces, $|I - I_n| < \varepsilon$. Sea $M(\varepsilon) = \sup\{K(\varepsilon), a + N(\varepsilon)\} + 1$ y sea $c > M(\varepsilon)$. Entonces, existe un número natural $n \geq N(\varepsilon)$ tal que $K(\varepsilon) \leq a + n < c$. Por lo tanto, la integral parcial I_c está dada por

$$I_c = \int_a^c f = \int_a^{a+n} f + \int_{a+n}^c f,$$

por lo que se deduce que $|I - I_c| < 2\varepsilon$. Q.E.D.

En el caso en que $f(x) \geq 0$ para toda $x \geq a$, el siguiente resultado proporciona una prueba útil.

32.6 TEOREMA *Supóngase que* $f(x) \geq 0$ *para toda* $x \geq a$ *y que* f *es integrable sobre* $[a, c]$ *para toda* $c \geq a$. *Entonces, la integral infinita de* f *existe si y sólo si el conjunto* $\{I_c : c \leq a\}$ *es acotado. En este caso*

$$\int_a^{+\infty} f = \sup\left\{\int_a^c f : c \geq a\right\}.$$

DEMOSTRACION. La hipótesis de que $f(x) \geq 0$ implica que I_c es una función monótonamente creciente de c. Por lo tanto, la existencia del lim xi es equivalente a la acotación de $\{I_c : c \geq a\}$. Q.E.D.

32.7 PRUEBA DE COMPARACION. *Suponga que* f *y* g *son integrables sobre* $[a, c]$ *para toda* $c \geq a$ *y que* $|f(x)| \leq g(x)$ *para toda* $x \geq a$. *Si la integral infinita de* g *existe, entonces, la integral infinita de* f *existe y*

$$\left| \int_a^{+\infty} f \right| \leq \int_a^{+\infty} g.$$

DEMOSTRACION. Si $a \leq c < b$, entonces, del lema 30.5 se sigue que $|f|$ es integrable en $[c, b]$ y que

$$\left| \int_c^b f \right| \leq \int_c^b |f| \leq \int_c^b g.$$

Se sigue del criterio de Cauchy 32.5 que las integrales infinitas de f y $|f|$ existen. Más aún, se tienen

$$\left| \int_a^{+\infty} f \right| \leq \int_a^{+\infty} |f| \leq \int_a^{+\infty} g.$$ Q.E.D.

32.8 PRUEBA DE COMPARACION DEL LIMITE. *Suponga que f y g son no negativas e integrables sobre $[a, c]$ para toda $c \geq a$ y que*

(32.12) $$\lim_{x \to +\infty} \frac{f(x)}{g(x)} \neq 0.$$

Entonces, las dos o ninguna de las integrales infinitas $\int_a^{+\infty} f, \int_a^{+\infty} g$ existen.

DEMOSTRACION. Por la relación (32.12) se deduce que existen números positivos $A < B$ y $K \geq a$ tales que

$$Ag(x) \leq f(x) \leq Bg(x) \quad \text{para } x \geq K.$$

La prueba de comparación 32.7 y esta relación muestran que ambas o ninguna de las integrales infinitas $\int_K^{+\infty} f, \int_K^{+\infty} g$ existen. Dado que, tanto f como g son integrables en $[a, K]$, la afirmación se sigue. Q.E.D.

32.9 PRUEBA DE DIRICHLET. *Suponga que f es continua para $x \geq a$, que las integrales parciales*

$$I_c = \int_a^c f, \qquad c \geq a,$$

son acotadas y que φ es monótonamente decreciente a cero conforme $x \to +\infty$. Entonces, la integral infinita $\int_a^{+\infty} f\varphi$ existe.

DEMOSTRACION. Sea A cota para el conjunto $\{|I_c| : c \geq a\}$. Si $\varepsilon > 0$, sea $K(\varepsilon)$ tal que si $x \geq K(\varepsilon)$, entonces $0 \leq \varphi(x) \leq \varepsilon/2A$. Si $b \geq c \geq K(\varepsilon)$, entonces, del ejercicio 30.N se sigue que existe un número ξ en $[c, b]$ tal que

$$\int_c^b f\varphi = \varphi(c) \int_c^{\xi} f.$$

A partir del cálculo

$$\left| \int_c^{\xi} f \right| = |I_{\xi} - I_c| \leq 2A,$$

se tiene que

$$\left| \int_c^b f\varphi \right| < \varepsilon$$

en donde $b \geq c$ exceden a $K(\varepsilon)$. Se puede aplicar entonces el criterio de Cauchy 32.5. Q.E.D.

32.10 EJEMPLOS. *(a)* Si $f(x) = 1/(1 + x^2)$ y $g(x) = 1/x^2$ para $x \geq a > 0$, entonces, $0 \leq f(x) \leq g(x)$. Como ya se ha visto en el ejemplo

32.4*(b)* que la integral infinita $\int_1^{+\infty} (1/x^2)\, dx$ existe, se sigue la prueba de comparación 32.7 que la integral infinita $\int_1^{+\infty} (1/(1+x^2))\, dx$ también existe. (Esto se podría probar directamente observando que

$$\int_1^c \frac{1}{1+x^2}\, dx = \text{Arc tan } c - \text{Arc tan } 1$$

y que Arc tan $c \to \pi/2$ conforme $c \to +\infty$.)

(b) Si $h(x) = e^{-x^2}$ y $g(x) = e^{-x}$ entonces, $0 \le h(x) \le g(x)$ para $x \ge 1$. En el ejemplo 32.4*(c)* se vio que la integral infinita $\int_0^{+\infty} e^{-x}\, dx$ existe, por lo tanto, se infiere de la prueba de comparación 32.7 que la integral infinita $\int_0^{+\infty} e^{-x^2}\, dx$ también existe. En este caso, no es posible un cálculo directo de las integrales parciales usando funciones elementales. Sin embargo, más adelante se verá que esta integral infinita es igual a $\frac{1}{2}\sqrt{\pi}$.

(c) Sea $p > 0$ y considere la existencia de la integral infinita

$$\int_1^{+\infty} \frac{\text{sen}\, x}{x^p}\, dx.$$

Si $p > 1$, el integrando está dominado por $1/x^p$, que se vio que es convergente en el ejemplo 3.4*(b)*. En este caso la prueba de comparación implica que la integral infinita converge. Si $0 < p \le 1$, este argumento no es válido; sin embargo, si se fijan $f(x) = \text{sen}\, x$ y $\varphi(x) = 1/x^p$, entonces, la prueba de Dirichlet 32.9 demuestra que la integral infinita existe.

(d) Sea $f(x) = \text{sen}\, x^2$ para $x \ge 0$ y considere la **integral de Fresnel**†

$$\int_0^{+\infty} \text{sen}\, x^2\, dx.$$

Es claro que la integral sobre $[0, 1]$ existe, de modo que se analizará solamente la integral sobre $\{x : x \ge 1\}$. Si se hace la substitución $t = x^2$ y se aplica el teorema del cambio de variable 30.12, se obtiene

$$\int_1^c \text{sen}\, x^2\, dx = \frac{1}{2}\int_1^{c^2} \frac{\text{sen}\, t}{\sqrt{t}}\, dt.$$

El ejemplo anterior muestra que la integral de la derecha converge cuando $c \to +\infty$; por lo que se sigue que $\int_1^{+\infty} \text{sen}\, x^2\, dx$ existe. (Observe que el integrando no converge a 0 conforme $x \to +\infty$.)

(e) Suponga que $\alpha \ge 1$ y defina $\Gamma(\alpha)$ por medio de la integral

$$\Gamma(\alpha) = \int_0^{+\infty} e^{-x}x^{\alpha-1}\, dx. \qquad (32.13)$$

Para poder ver que esta integral infinita existe, considere la función $g(x) = 1/x^2$ para $x \ge 1$. Dado que

† AUGUSTIN FRESNEL (1788-1827) fue un físico matemático francés, ayudó a restablecer la teoría ondulatoria de la luz que se introdujo anteriormente por Huygens.

$$\lim_{x\to+\infty} \frac{e^{-x}x^{\alpha-1}}{x^{-2}} = \lim_{x\to+\infty} \frac{x^{\alpha+1}}{e^{x}} = 0,$$

se tiene que si $\varepsilon > 0$ entoncei:, existe $K(\varepsilon)$ tal que

$$0 < e^{-x}x^{\alpha-1} \leq \varepsilon x^{-2} \quad \text{para } x \leq K(\varepsilon).$$

Como la integral infinita $\int_K^{+\infty} x^{-2}\, dx$ existe, se deduce que la integral (32.13) también converge. La función definida para $\alpha \geq 1$ por medio de la fórmula (32.13) es conocida como **función gama.** Rápidamente se verá que si $\alpha < 1$, entonces, el integrando $e^{-x}x^{\alpha-1}$ llega a ser no acotado cerca de $x = 0$. Sin embargo, si α satisface $0 < \alpha < 1$, en el ejemplo 32.2*(c)* se ha visto que la función $x^{\alpha-1}$ tiene una integral impropia sobre el intervalo $[0, 1]$. Dado que $0 < e^{-x} \leq 1$ para toda $x \geq 0$, fácilmente se prueba que la integral impropia

$$\int_{0+}^{1} e^{-x}x^{\alpha-1}\, dx$$

existe cuando $0 < \alpha < 1$. Por lo tanto, se puede extender la definición de la función gama que se dé para toda $\alpha > 0$ por una integral de la forma (32.13) siempre que se interprete como una suma

$$\int_{0+}^{a} e^{-x}x^{\alpha-1}\, dx + \int_{a}^{+\infty} e^{-x}x^{\alpha-1}\, dx$$

de una integral impropia y de una integral infinita.

Convergencia absoluta

Si f es Riemann integrable en $[a, c]$ para todas $c \geq a$, entonces, se sigue del teorema 30.4*(a)* que $|f|$, el valor absoluto de f, también es Riemann integrable en $[a, c]$ para $c \geq a$. De la prueba de comparación 32.7 se sigue que si la integral infinita

(32.14) $$\int_a^{+\infty} |f(x)|\, dx$$

existe, entonces, la integral infinita

(32.15) $$\int_a^{+\infty} f(x)\, dx$$

también existe y está acotada en valor absoluto por (32.14)

31.11 DEFINICION. Si la integral infinita (32 14) existe, entonces, se dice que f es **absolutamente integrable** sobre $\{x : x \geq a\}$, o que la integral infinita (32.15) es **absolutamente convergente.**

Se ha insistido en que si f es absolutamente integrable sobre $\{x : x \geq a\}$, entonces, la integral infinita (32.15) existe. Sin embargo, lo inverso no es cierto, como se puede ver al considerar la integral

$$\int_{\pi}^{+\infty} \frac{\operatorname{sen} x}{x}\, dx.$$

La convergencia de esta integral se probó en el ejemplo 32.10*(c)*. No obstante, se puede ver fácilmente que en cada intervalo $[k\pi, (k+1)\pi]$, $k \in \mathbf{N}$, hay un subintervalo de longitud $b > 0$ en el que

$$|\operatorname{sen} x| \geq \tfrac{1}{2}.$$

(De hecho, se puede tomar $b = 2\pi/3$.) Por lo tanto, se tiene

$$\int_{\pi}^{k\pi} \left|\frac{\operatorname{sen} x}{x}\right| dx = \int_{\pi}^{2\pi} + \cdots + \int_{(k-1)\pi}^{k\pi} \geq \frac{b}{2}\left\{\frac{1}{2\pi} + \frac{1}{3\pi} + \cdots + \frac{1}{k\pi}\right\},$$

y se sigue (ver 16.11*(c)*) que la función $f(x) = \operatorname{sen} x/x$ no es absolutamente integrable sobre $\{x : x \geq \pi\}$.

Se puede ver que la prueba de comparación 32.7, de hecho, prueba la **convergencia absoluta** de la integral infinita de f sobre el intervalo $[a, +\infty)$.

Ejercicios

32.A. Suponga que f es una función acotada de valor real en $J = [a, b]$ y que f es integrable sobre $[c, b]$ para toda $c > a$. Demostrar que la integral impropia de f sobre J existe.

32.B. Suponga que f es integrable sobre $[c, b]$ para toda $c > a$ y que la integral impropia $\int_{a+}^{b} |f|$ existe. demostrar que la integral impropia $\int_{a+}^{b} f$ existe, pero que lo inverso puede no ser cierto.

32.C. Suponga que f y g son integrables en $[c, b]$ para toda $c \in (a, b)$. Si $|f(x)| \leq g(x)$ para $x \in J = [a, b]$ y si g tiene una integral impropia en J, entonces, también f la tiene.

32.D. Analizar la convergencia o la divergencia de las siguientes integrales impropias:

(a) $\displaystyle\int_0^1 \frac{dx}{(x + x^2)^{1/2}}$, (b) $\displaystyle\int_0^1 \frac{dx}{(x - x^2)^{1/2}}$,

(c) $\displaystyle\int_0^1 \frac{x\, dx}{(1 - x^3)}$, (d) $\displaystyle\int_0^1 \frac{\log x}{\sqrt{x}}\, dx$,

(e) $\displaystyle\int_0^1 \frac{\log x}{1 - x^2}\, dx$, (f) $\displaystyle\int_0^1 \frac{x\, dx}{(1 - x^3)^{1/2}}$.

32.E. Determinar los valores de p y q para los cuales las siguientes integrales convergen:

(a) $\displaystyle\int_0^1 x^p (1 - x)^q\, dx$, (b) $\displaystyle\int_0^{\pi/2} x^p (\operatorname{sen} x)^q\, dx$,

(c) $\displaystyle\int_1^2 (\log x)^p\, dx$, (d) $\displaystyle\int_0^1 x^p (-\log x)^q\, dx$.

32.F. Analizar la convergencia o la divergencia de las siguientes integrales. ¿Cuáles son absolutamente convergentes?

(a) $\displaystyle\int_1^{+\infty} \frac{dx}{x(1+\sqrt{x})}$, (b) $\displaystyle\int_1^{+\infty} \frac{x+2}{x^2+1}\, dx$,

(c) $\displaystyle\int_1^{+\infty} \frac{\operatorname{sen}(1/x)}{x}\, dx$, (d) $\displaystyle\int_1^{+\infty} \frac{\cos x}{\sqrt{x}}\, dx$,

(e) $\displaystyle\int_0^{+\infty} \frac{x \operatorname{sen} x}{1+x^2}\, dx$, (f) $\displaystyle\int_0^{+\infty} \frac{\operatorname{sen} x \operatorname{sen} 2x}{x}\, dx$.

32.G. ¿Para qué valores de p y q son convergentes las siguientes integrales? ¿para qué valores son absolutamente convergentes?

(a) $\displaystyle\int_1^{+\infty} \frac{x^p}{1+x^q}\, dx$, (b) $\displaystyle\int_1^{+\infty} \frac{\operatorname{sen} x}{x^q}\, dx$,

(c) $\displaystyle\int_1^{+\infty} \frac{\operatorname{sen} x^p}{x}\, dx$, (d) $\displaystyle\int_1^{+\infty} \frac{1-\cos x}{x^q}\, dx$.

32.H. Si f es integrable en cualquier intervalo $[0, c]$ para si la integral infinita $c > 0$ existe.

32.I. Dar un ejemplo en donde la integral infinita $\int_0^{+\infty} f$ exista pero que f no esté acotada en el conjunto $\{x : x \geq 0\}$.

32.J. Si f es monótona y la integral infinita $\int_0^{+\infty} f$ existe, entonces, $xf(x) \to 0$ conforme $x \to +\infty$.

Sección 33 Convergencia uniforme e integrales infinitas

En muchas aplicaciones es importante considerar integrales infinitas en las que el integrando dependa de un parámetro. Para poder manejar fácilmente esta situación el concepto de convergencia uniforme de la integral con respecto al parámetro es de fundamental importancia. Primero se considerará el caso en el que el parámetro pertenece a un intervalo $J = [\alpha, \beta]$.

33.1. DEFINICION. Sea f una función de valor real definida para (x, t) tal que satisfaga $x \geq a$ y $\alpha \leq t \leq \beta$. Suponga que para cada t en $J = [\alpha, \beta]$ la integral infinita

$$F(t) = \int_a^{+\infty} f(x, t)\, dx \tag{33.1}$$

existe. Se dice que esta convergencia es **uniforme en** J si para toda $\varepsilon > 0$ existe un número $N(\varepsilon)$ tal que si $c \geq N(\varepsilon)$ y $t \in J$, entonces,

$$\left| F(t) - \int_a^c f(x, t)\, dx \right| < \varepsilon.$$

La diferencia entre la convergencia ordinaria de las integrales infinitas dadas en (33.1) y la convergencia uniforme es que $M(\varepsilon)$ se puede elegir independientemente del valor de t en J. Se le deja al lector escribir la definición de convergencia uniforme de las integrales infinitas cuando el parámetro t pertenece al conjunto $\{t : t \geq \alpha\}$ o al conjunto N.

Es útil tener algunas pruebas para convergencia uniforme de las integrales infinitas.

33.2 CRITERIO DE CAUCHY. *Suponga que para cada $t \in J$, la integral infinita (33.1) existe. Entonces, la convergencia es uniforme en J si sólo si para cada $\varepsilon > 0$ existe un número $K(\varepsilon)$ tal que si $b \geq c \geq K(\varepsilon)$ y $t \in J$, entonces,*

$$\left| \int_c^b f(x, t)\, dx \right| < \varepsilon. \tag{33.2}$$

La demostración queda como ejercicio.

33.3 PRUEBA M DE WEIERSTRASS. *Suponga que f es Riemann integrable sobre $[a, c]$ para todas $c \geq a$ y toda $t \in J$. Suponga que existe una función positiva M definida para $x \geq a$ tal que*

$$|f(x, t)| \leq M(x) \quad \text{para } x \geq a,\ t \in J,$$

y tal que la integral infinita $\int_a^{+\infty} M(x)\, dx$ exista. Entonces, para cada $t \in J$, la integral en (33.1) es (absolutamente) convergente y la convergencia es uniforme en J.

DEMOSTRACION. La convergencia de

$$\int_a^{+\infty} |f(x, t)|\, dx \quad \text{para } t \in J,$$

es consecuencia inmediata de la prueba de comparación y de las hipótesis. Por lo tanto la integral que da $F(t)$ es absolutamente convergente para $t \in J$. Usando el criterio de Cauchy junto con el hecho

$$\left| \int_c^b f(x, t)\, dx \right| \leq \int_c^b |f(x, t)|\, dx \leq \int_c^b M(x)\, dx,$$

fácilmente se puede probar la convergencia uniforme en J. Q.E.D.

La prueba M de Weierstrass es útil cuando la convergencia es absoluta asi como uniforme, pero no es lo suficientemente sutil para tratar el caso de convergencia no absolutamente uniforme. Para esto, se dará una analogía de la prueba de Dirichlet 32.9.

33.4 PRUEBA DE DIRICHLET. *Sea f continua en (x, t) para $x \geq a$ y t en J y suponga que existe una constante A tal que*

$$\left| \int_a^c f(x, t)\, dx \right| \le A \quad \text{para } c \ge a, \quad t \in J.$$

Suponga que para cada $t \in J$, *la función* $\varphi(x, t)$ *es monótonamente decreciente para* $x \ge a$, *y converge uniformemente a* 0 *conforme* $x \to +\infty$ *para* $t \in J$. *Entonces, la integral*

$$F(t) = \int_a^{+\infty} f(x, t)\varphi(x, t)\, dx$$

converge uniformemente en J.

DEMOSTRACION. Sea $\varepsilon > 0$ y elíjase $K(\varepsilon)$ tal que $x \ge K(\varepsilon)$ y $t \in J$, entonces, $\varphi(x, t) < \varepsilon/2A$. Si $b \ge c \ge K(\varepsilon)$, entonces, del ejercicio 30.N se sigue que para cada $t \in J$, existe un número $\xi(t)$ en $[c, b]$ tal que

$$\int_c^b f(x, t)\varphi(x, t)\, dx = \varphi(c, t) \int_c^{\xi(t)} f(x, t)\, dx.$$

Por lo tanto, si $b \ge c \ge K(\varepsilon)$ y $t \in J$, se tiene

$$\left| \int_c^b f(x, t)\varphi(x, t)\, dx \right| \le \varphi(c, t) 2A < \varepsilon,$$

de modo que la uniformidad de la convergencia se sigue del criterio de Cauchy 33.2.

Q.E.D.

33.5 EJEMPLOS. *(a)* Si f está dada por

$$f(x, t) = \frac{\cos tx}{1 + x^2}, \qquad x \ge 0, \quad t \in \mathbf{R},$$

y si se define $M(x) = (1 + x^2)^{-1}$, entonces $|f(x, t)| \le M(x)$. Dado que la integral infinita de M en $[0, +\infty)$ existe, de la prueba M de Weierstrass se tiene que la integral infinita

$$\int_0^{+\infty} \frac{\cos tx}{1 + x^2}\, dx$$

converge uniformemente para $t \in \mathbf{R}$.

(b) Sea $f(x, t) = e^{-x} x^t$ para $x \ge 0$, $t \ge 0$. Se puede ver que la integral

$$\int_0^{+\infty} e^{-x} x^t\, dx$$

converge uniformemente para t en un intervalo $[0, \beta]$ para cualquier $\beta > 0$. Sin embargo, no converge uniformemente en $\{t \in \mathbf{R} : t \ge 0\}$. (Ver el ejercicio 33.A.)

(c) Si $f(x, t) = e^{-tx}$ sen x para $x \ge 0$ y $t \ge \gamma > 0$, entonces,

$$|f(x, t)| \le e^{-tx} \le e^{-\gamma x}.$$

Si se fija $M(x)=e^{-\gamma x}$, entonces, la prueba M de Weierstrass implica que la integral

$$\int_0^{+\infty} e^{-tx} \operatorname{sen} x \, dx$$

converge uniformemente para $t \geq \gamma > 0$ y un cálculo elemental prueba que converge a $(1+t^2)^{-1}$. (Observe que si $t=0$, entonces, la integral ya no converge.)

(d) Considere la integral infinita

$$\int_0^{+\infty} e^{-tx} \frac{\operatorname{sen} x}{x} \, dx \qquad \text{para } t \geq 0,$$

en donde el integrando es 1 para $x=0$. Dado que el integrando está dominado por 1 es suficiente demostrar que la integral sobre $\varepsilon \leq x$ converge uniformemente para $t \geq 0$. La prueba M de Weierstrass no es aplicable a este integrando. Sin embargo, si se toman $f(x,t)=\operatorname{sen} x$ y $\varphi(x,t)=e^{-tx}/x$, entonces, las hipótesis de la prueba de Dirichlet se cumplen.

Integrales infinitas dependiendo de un parámetro

Suponga que f es una función continua de (x, t) definida para $x \geq a$ y para t en $J=[\alpha, \beta]$. Más aún, suponga que la integral infinita

$$F(t)=\int_a^{+\infty} f(x,t)\,dx \tag{33.1}$$

existe para cada $t \in J$. Se habrá de probar que si esta convergencia es uniforme entonces F es continúa en J y su integrando se puede calcular intercambiando el orden de integración. Se demostrará un resultado análogo para la derivada.

33.6 TEOREMA. *Suponga que f es continua en (x, t) para $x \geq a$ y t en $J=[\alpha, \beta]$ y que la convergencia en (33.1) es uniforme en J. Entonces, F es continua en J.*

DEMOSTRACION. Si $n \in \mathbf{N}$, defina F_n en J por medio de

$$F_n(t)=\int_a^{a+n} f(x,t)\,dx.$$

Del teorema 31.6 se tiene que F_n es continua en J. Dado que la sucesión (F_n) converge a F uniformemente en J, del teorema 24.1 se sigue que F es continua en J. Q.E.D.

33.7 TEOREMA. *Bajo las hipótesis del teorema anterior se tiene*

$$\int_{\alpha}^{\beta} F(t)\,dt = \int_{a}^{+\infty}\left\{\int_{\alpha}^{\beta} f(x, t)\,dt\right\} dx,$$

que se puede escribir de la forma

$$\int_{\alpha}^{\beta}\left\{\int_{a}^{+\infty} f(x, t)\,dx\right\} dt = \int_{a}^{+\infty}\left\{\int_{\alpha}^{\beta} f(x, t)\,dt\right\} dx. \qquad (33.3)$$

DEMOSTRACION. Si F_n está definida como en la demostración anterior, entonces, del teorema 31.9 se sigue que

$$\int_{\alpha}^{\beta} F_n(t)\,dt = \int_{a}^{a+n}\left\{\int_{\alpha}^{\beta} f(x, t)\,dt\right\} dx.$$

Dado que (F_n) converge a F uniformemente en J, entonces, el teorema 31.2 implica que

$$\int_{\alpha}^{\beta} F(t)\,dt = \lim_{n}\int_{\alpha}^{\beta} F_n(t)\,dt.$$

Combinando las dos últimas relaciones, se obtiene (33.3). Q.E.D.

33.8 TEOREMA. *Suponga que f y su derivada parcial f_t son continuas en (x, t) para $x \geq a$ y t en $J = [\alpha, \beta]$. Suponga que (33.1) existe para toda $t \in J$ y que*

$$G(t) = \int_{a}^{+\infty} f_t(x, t)\,dx$$

es uniformemente convergente en J. Entonces, F es diferenciable en J y $F' = G$. En símbolos:

$$\frac{d}{dt}\int_{a}^{+\infty} f(x, t)\,dx = \int_{a}^{+\infty} \frac{\partial f}{\partial t}(x, t)\,dx.$$

DEMOSTRACION. Si F_n est definida para $t \in J$ como

$$F_n(t) = \int_{a}^{a+n} f(x, t)\,dx,$$

entonces, del teorema 31.7 se sigue que F_n es diferenciable y que

$$F_n'(t) = \int_{a}^{a+n} f_t(x, t)\,dx.$$

Por hipótesis, la sucesión (F_n) converge en J a F y la sucesión (F_n') converge uniformemente en J a G. Del teorema 28.5 se sigue que F es diferenciable en J y que $F' = G$. Q.E.D.

33.9 EJEMPLOS. *(a)* Se observa que si $t > 0$, entonces

$$\frac{1}{t} = \int_0^{+\infty} e^{-tx}\, dx$$

y que la convergencia es uniforme para $t \geq t_0 > 0$. Si se integran los dos lados de esta relación con respecto a t sobre un intervalo $[\alpha, \beta]$ en donde $0 < \alpha < \beta$, y se usa el teorema 33.7, si obtiene la fórmula

$$\log(\beta/\alpha) = \int_\alpha^\beta \frac{1}{t}\, dt = \int_0^{+\infty} \left\{ \int_\alpha^\beta e^{-tx}\, dt \right\} dx$$
$$= \int_0^{+\infty} \frac{e^{-\alpha x} - e^{-\beta x}}{x}\, dx.$$

(Observe que el último integrando se puede definir como continuo en $x = 0$.)

(b) En lugar de integrar con respecto a t se diferencia y se obtiene formalmente

$$\frac{1}{t^2} = \int_0^{+\infty} x e^{-tx}\, dx.$$

Dado que esta última integral converge uniformemente con respecto a t, siempre que $t \geq t_0 > 0$, la fórmula es válida para $t > 0$. Por inducción se obtiene

$$\frac{n!}{t^{n+1}} = \int_0^{+\infty} x^n e^{-tx}\, dx \quad \text{para } t > 0.$$

Haciendo referencia a la definición de la función gama dada en el ejemplo 32.10 (e) se puede ver que $\Gamma(n+1) = n!$.

(c) Si $\alpha > 1$ es un número real y $x > 0$, entonces, $x^{\alpha-1} = e^{(\alpha-1)\log x}$. Por lo que $f(\alpha) = x^{\alpha-1}$ es un función continua de (α, x). Más aún, se puede ver que existe una vecindad de α en la que la integral

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x}\, dx$$

es uniformemente convergente. Del teorema 33.6 se sigue que la función gama es continua cuando menos para $\alpha > 1$. (Si $0 < \alpha \leq 1$, se puede obtener la misma conclusión, pero, se debe tomar en cuenta el hecho de que la integral es impropia en $x = 0$)

(d) Sean $t \geq 0$ $u \geq 0$ y definase a F por medio de

$$F(u) = \int_0^{+\infty} e^{-tx} \frac{\operatorname{sen} ux}{x}\, dx.$$

Si $t > 0$, entonces, esta integral es uniformemente convergente para $u \geq 0$ lo mismo que la integral

$$F'(u) = \int_0^{+\infty} e^{-tx} \cos ux\, dx.$$

Más aún, la integración por partes muestra que

$$\int_0^A e^{-tx} \cos ux\, dx = \left[\frac{e^{-tx}[u \operatorname{sen} ux - t \cos ux]}{t^2+u^2}\right]_{x=0}^{x=A}.$$

Si se toma $A \to +\infty$, se obtiene la fórmula

$$F'(u) = \int_0^{+\infty} e^{-tx} \cos ux\, dx = \frac{t}{t^2+u^2}, \qquad u \geq 0.$$

Por lo tanto, existe una constante C tal que

$$F(u) = \operatorname{Arc\,tan}(u/t) + C \qquad \text{para} \qquad u \geq 0.$$

Para poder calcular la constante C, se usa el hecho de que $F(0) = 0$ y Arc tan $(0) = 0$ y se deduce que $C = 0$. Por lo tanto tanto, si $t > 0$ y $u \geq 0$, entonces

$$\operatorname{Arc\,tan}(u/t) = \int_0^{+\infty} e^{-tx} \frac{\operatorname{sen} ux}{x}\, dx.$$

(e) Ahora, mantenga $u > 0$ fija en la última fórmula y observe, como en el ejemplo 33.5 (d), que la integral converge uniformemente para $t \geq 0$ de tal manera que el límite es continuo para $t \geq 0$. Si se toma $t \to 0+$, se obtiene la importante fórmula

(33.4) $$\frac{\pi}{2} = \int_0^{+\infty} \frac{\operatorname{sen} ux}{x}\, dx, \qquad u > 0.$$

Integrales infinitas de sucesiones

Sea (f_n) una sucesión de funciones de valor real definidas para $x \geq a$ se habrá de suponer que las integrales infinitas $\int_a^{+\infty} f_n$ existen y que el límite $f(x) = \lim (f_n(x))$ existe para cada $x \geq a$. Convendría poder concluir que la integral infinita de f existe y que

(33.5) $$\int_a^{+\infty} f = \lim \int_a^{+\infty} f_n.$$

En el teorema 31.2 se demostró que si una sucesión (f_n) de funciones Riemann integrables converge uniformemente en un intervalo $[a, c]$ a una función f, entonces, f es Riemann integrable y la integral de f es el límite de las integrales de las f_n. El resultado correspondiente no necesariamente es verdadero para integrales infinitas; en el ejercicio 33.J se verá que la función límite no tiene que poseer una integral infinita. Además, aun cuando la integral infinita exista y ambos lados de (33.5) tengan un significado, la igualdad puede no cumplirse (cf. ejercicio 33.K) Análogamente la extensión inmediata del teorema de convergencia acotada 31.3 puede no ser válida para integrales infinitas. Sin embargo, existen dos resultados importantes y útiles que dan las condiciones bajo las cuales la ecuación (33.5) se cumple. Al demostrarlas se

hará uso del teorema de convergencia acotada 31.3. El primer resultado es un caso especial del célebre teorema atribuido a Lebesgue. (Dado que estamos tratando con integrales infinitas de Riemann, es necesario agregar la hipótesis de que la función límite es integrable. En la teoría (más general) de integración de Lebesgue, esta hipótesis adicional no es necesaria)

33.10 TEOREMA DE CONVERGENCIA DOMINADA. *Suponga que* (f_n) *es una sucesión acotada de funciones de valor real, que* $f(x) = \lim (f_n(x))$ *para toda* $x \geq a$, *y que* f *y* f_n, $n \in \mathbf{N}$, *son Riemann integrables sobre* $[a, c]$ *para toda* $c > a$. *Suponga que existe una función* M *que tiene una integral sobre* $x \geq a$ *y que*

$$|f_n(x)| \leq M(x) \quad \text{para } x \geq a, \quad n \in \mathbf{N}.$$

Entonces, f tiene una integral sobre $x \geq a$ y

$$\int_a^{+\infty} f = \lim \int_a^{+\infty} f_n. \tag{33.5}$$

DEMOSTRACION. De la prueba de comparación 32.7 se sigue que las integrales infinitas

$$\int_a^{+\infty} f, \qquad \int_a^{+\infty} f_n, \qquad n \in \mathbf{N},$$

existen. Si $\varepsilon > 0$, sea K tal que $\int_K^{+\infty} M < \varepsilon$, de donde se sigue que

$$\left| \int_K^{+\infty} f \right| < \varepsilon \qquad \text{y} \qquad \left| \int_K^{+\infty} f_n \right| < \varepsilon, \qquad n \in \mathbf{N}.$$

Dado que $f(x) = \lim (f_n(x))$ para toda $x \in [a, K]$ se sigue del teorema de convergencia acotada 31.3 que $\int_a^K f = \lim_n \int_a^K f_n$. Por lo tanto, tiene

$$\left| \int_a^{+\infty} f - \int_a^{+\infty} f_n \right| \leq \left| \int_a^K f - \int_a^K f_n \right| + 2\varepsilon,$$

que es menor que 3ε para n suficientemente grande. Q.E.D.

33.11 TEOREMA DE CONVERGENCIA MONOTONA. *Supongase que* (f_n) *es una sucesión acotada de funciones positivas en* $\{x : a \geq a\}$ *que es monótonamente creciente en el sentido de que* $f_n(x) \leq f_{n+1}(x)$ *para* $n \in \mathbf{N}$ *y* $x \geq a$, *y tal que* f *y cada* f_n *tiene una integral sobre* $[a, c]$ *para toda* $c > a$. *Entonces, a función límite* f *tiene una integral sobre* $\{x : x \geq a\}$ *sí y sólo si el conjunto* $\{\int_a^{+\infty} f_n : n \in \mathbf{N}\}$ *es acotado. En este caso*

$$\int_a^{+\infty} f = \sup_n \left\{ \int_a^{+\infty} f_n \right\} = \lim_n \int_a^{+\infty} f_n.$$

DEMOSTRACION. Dado que la sucesión (f_n) es monótonamente creciente, se deduce que la sucesión $(\int_a^{+\infty} f_n : n \in \mathbf{N})$ también es monótona-

mente creciente. Si f tiene una integral sobre $\{x : x \geq a\}$, entonces, el teorema de convergencia dominada (con $M = f$) prueba que

$$\int_a^{+\infty} f = \lim \int_a^{+\infty} f_n.$$

Inversamente, suponga que el conjunto de integrales infinitas es acotado y sea S el supremo de este conjunto. Si $c > a$, entonces, el teorema de convergencia monótona 31.4 implica que

$$\int_a^c f = \lim_n \int_a^c f_n = \sup_n \left\{ \int_a^c f_n \right\}.$$

Dado que $f_n \geq 0$, se sigue que $\int_a^c f_n \leq \int_a^{+\infty} f_n \leq S$, y por lo tanto, que $\int_a^c f \leq S$. Por el teorema 32.6 la integral infinita de f existe y

$$\int_a^{+\infty} f = \sup_c \int_a^c f = \sup_c \left\{ \sup_n \int_a^c f_n \right\}$$

$$= \sup_n \left\{ \sup_c \int_a^c f_n \right\} = \sup_n \int_a^{+\infty} f_n. \qquad \text{Q.E.D.}$$

Integrales iteradas e infinitas

En el teorema 33.7 se obtuvo un resultado que justifica el intercambio del orden de integración sobre la región $\{(x, t) : a \leq x, \alpha \leq t \leq \beta\}$. Tambien es deseable poder intercambiar ei orden de integración de una integral iterada infinita. Es decir, se desea probar la igualdad

$$\text{(33.6)} \qquad \int_\alpha^{+\infty} \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt = \int_a^{+\infty} \left\{ \int_\alpha^{+\infty} f(x, t)\, dt \right\} dx,$$

condición sencilla que también implique convergencia absoluta de las integrales. Sin embargo, para poder estudiar integrales iteradas infinitas, que no necesariamente sean absolutamente convergentes, se requiere de una serie de condiciones más complicadas.

33.12 TEOREMA: *Suponga que f es una función positiva definida para (x, t) que satisface $x \geq a$, $t \geq \alpha$. Suponga que*

$$\text{(33.7)} \qquad \int_a^b \left\{ \int_\alpha^{+\infty} f(x, t)\, dt \right\} dx = \int_\alpha^{+\infty} \left\{ \int_a^b f(x, t)\, dx \right\} dt$$

para cada $b \geq a$ y que

$$\text{(33.7')} \qquad \int_\alpha^\beta \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt = \int_a^{+\infty} \left\{ \int_\alpha^\beta f(x, t)\, dt \right\} dx$$

para cada $\beta \geq \alpha$. Entonces, si una de las integrales iteradas de la ecuación (33.6) existe, las otras también existen y son iguales.

DEMOSTRACION. Suponga que la integral del lado izquierdo en (33.6) existe. Dado que f es positiva

$$\int_a^b f(x, t)\, dx \le \int_a^{+\infty} f(x, t)\, dx$$

para cada $b \ge a$ y $t \ge \alpha$. Por lo tanto, se infiere de la prueba de comparación 32.7 que

$$\int_\alpha^{+\infty} \left\{ \int_a^b f(x, t)\, dx \right\} dt \le \int_\alpha^{+\infty} \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt.$$

Usando la relación (33.7), se concluye que

$$\int_a^b \left\{ \int_\alpha^{+\infty} f(x, t)\, dt \right\} dx \le \int_\alpha^{+\infty} \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt$$

para cada $b \ge a$. Una aplicación del teorema 32.6 muestra que se puede tomar el límite cuando $b \to +\infty$, de tal manera que la otra integral iterada existe y

$$\int_a^{+\infty} \left\{ \int_\alpha^{+\infty} f(x, t)\, dt \right\} dx \le \int_\alpha^{+\infty} \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt.$$

Si se repite este argumento y se aplica la ecuación (33.7'), se obtiene la desigualdad inversa. Por lo tanto, la igualdad se debe cumplir. Q.E.D.

33.13 TEOREMA. *Suponga que f es continua para $x \ge a$, $t \ge \alpha$, y que existen funciones positivas M y N tales que las integrales infinitas $\int_a^{+\infty} M$ y $\int_\alpha^{+\infty} N$ existen. Si la desigualdad*

(33.8) $$|f(x, t)| \le M(x)N(t), \qquad x \ge a, \quad t \ge \alpha,$$

se cumple, entonces, ambas integrales iteradas en (33.6) existen y son iguales.

DEMOSTRACION. Defina g para $x \ge a$, $t \ge \alpha$ por medio de $g(x, t) = f(x, t) + M(x)N(t)$ de tal manera que

$$0 \le g(x, t) \le 2M(x)N(t).$$

Dado que N es acotada en cada intervalo $[\alpha, \beta]$, se tiene de la desigualdad (33.8) y la prueba M de Weierstrass 33.3 que la integral

$$\int_a^{+\infty} g(x, t)\, dx$$

existe uniformemente para $t \in [\alpha, \beta]$ Aplicando el teorema 33.7 se puede ver que la ecuación (33.7') es válida (con f reemplazada por g) para cada $\beta \ge \alpha$. Análogamente, (33.7) es válida (con f reemplazada por g) para cada $b \ge a$.

También, la prueba de comparación 32.7 implica que las integrales en (33.6) existen (con f reemplazada por g). Del teorema 33.12 se deduce que estas integrales iteradas de g son iguales. Pero esto implica que las integrales iteradas de f existen y son iguales. Q.E.D.

Los resultados anteriores tratan el caso en que las integrales iteradas son absolutamente convergentes. En seguida se da un resultado en el que se trata el caso de convergencia no absoluta.

33.14 TEOREMA. *Suponga que la función de valor real f es continua en (x, t) para $x \geq a$ y $t \geq \alpha$ y que las integrales infinitas*

(33.9)
$$\int_a^{+\infty} f(x, t)\, dx, \qquad \int_\alpha^{+\infty} f(x, t)\, dt$$

convergen uniformemente para $t \geq \alpha$ y $x \geq a$, respectivamente. Además, defina a F para $x \geq a$, $\beta \geq \alpha$, por medio de

$$F(x, \beta) = \int_\alpha^{\beta} f(x, t)\, dt$$

y suponga que la integral infinita (33.10)

(33.10)
$$\int_a^{+\infty} F(x, \beta)\, dx$$

coverge uniformemente para $\beta \geq \alpha$. Entonces, ambas integrales iteradas infinitas existe y son iguales.

DEMOSTRACION. Como la integral infinita (33.10) es uniformemente convergente para $\beta \geq \alpha$, si $\varepsilon > 0$ existe un número $A_\varepsilon \geq a$ tal que si $A \geq A_\varepsilon$, entonces

(33.11)
$$\left| \int_a^{A} F(x, \beta)\, dx - \int_a^{+\infty} F(x, \beta)\, dx \right| < \varepsilon$$

para toda $\beta \geq \alpha$. También se puede ver que

$$\int_a^{A} F(x, \beta)\, dx = \int_a^{A} \left\{ \int_\alpha^{\beta} f(x, t)\, dt \right\} dx$$
$$= \int_\alpha^{\beta} \left\{ \int_a^{A} f(x, t)\, dx \right\} dt.$$

Por el teorema 33.7 y la convergencia uniforme de la segunda integral en (33.9) se deduce que

$$\lim_{\beta \to +\infty} \int_a^{A} F(x, \beta)\, dx = \int_a^{A} \left\{ \int_\alpha^{+\infty} f(x, t)\, dt \right\} dx.$$

Por lo que existe un número $B \geq \alpha$ tal que si $\beta_2 \geq \beta_1 \geq B$, entonces,

$$(33.12) \qquad \left| \int_a^A F(x, \beta_2)\, dx - \int_a^A F(x, \beta_1)\, dx \right| < \varepsilon.$$

Combinando (33.11) y (33.12) se puede ver que si $\beta_2 \geq \beta_1 \geq B$, entonces,

$$\left| \int_a^{+\infty} F(x, \beta_2)\, dx - \int_a^{+\infty} F(x, \beta_1)\, dx \right| < 3\varepsilon,$$

por lo que se tiene que al límite de $\int_a^{+\infty} F(x, \beta)\, dx$ existe cuando $\beta \to +\infty$. Después de aplicar el teorema 33.7 a la convergencia uniforme de la primera integral en (33.9), se tiene

$$\begin{aligned}\lim_{\beta\to+\infty} \int_a^{+\infty} F(x, \beta)\, dx &= \lim_{\beta\to+\infty} \int_a^{+\infty} \left\{ \int_\alpha^\beta f(x, t)\, dt \right\} dx \\ &= \lim_{\beta\to+\infty} \int_\alpha^\beta \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt \\ &= \int_\alpha^{+\infty} \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt.\end{aligned}$$

Dado que los dos términos del lado izquierdo de (33.11) tienen límites conforme $\beta \to +\infty$ se concluye, al pasar al límite, que

$$\left| \int_a^A \left\{ \int_\alpha^{+\infty} f(x, t)\, dt \right\} dx - \int_\alpha^{+\infty} \left\{ \int_a^{+\infty} f(x, t)\, dx \right\} dt \right| \leq \varepsilon.$$

Si se hace $A \to +\infty$, se obtiene la igualdad de las integrales iteradas impropias. Q.E.D.

Los teoremas recién dados que justifican el intercambio del orden de integración con frecuencia son útiles, pero aún dejan amplio margen para la inventiva. Frecuentemente se usa en forma conjunta con los teoremas 33.10 y 33.11 de convergencia dominada o monótona.

33.15 EJEMPLOS. *(a)* Si $f(x, t) = e^{-(x+t)}$ sen xt, entonces, se toma $M(x) = e^{-x}$ y $N(t) = e^{-t}$ para aplicar el teorema 33.13 y deducir que

$$\int_0^{+\infty} \left\{ \int_0^{+\infty} e^{-(x+t)} \operatorname{sen} xt\, dx \right\} dt = \int_0^{+\infty} \left\{ \int_0^{+\infty} e^{-(x+t)} \operatorname{sen} xt\, dt \right\} dx.$$

(b) Si $g(x, t) = e^{-xt}$, para $x \geq 0$ y $t \geq 0$, entonces, existe un problema en las rectas $x = 0$ y $t = 0$. Sin embargo, si $a > 0$, $\alpha > 0$, y $x \geq a$ y $t \geq \alpha$, se observa que

$$e^{-xt} = e^{-xt/2} e^{-xt/2} \leq e^{-\alpha x/2} e^{-at/2},$$

Si se fijan $M(x) = e^{-\alpha x/2}$ y $N(t) = e^{-at/2}$, entonces, el teorema 33.13 implica que

$$\int_a^{+\infty}\left\{\int_a^{+\infty} e^{-xt}\,dx\right\}dt = \int_a^{+\infty}\left\{\int_a^{+\infty} e^{-xt}\,dt\right\}dx.$$

(c) Considere la función $f(x, y) = xe^{-x^2(1+y^2)}$ para $x \geq a > 0$ $y \geq 0$. Si se toma $M(x) = xe^{-x^2}$ y $N(y) = e^{-a^2y^2}$, entonces, se puede invertir el orden de integración sobre $a \leq x$ y $0 \leq y$. Dado que se tiene

$$\int_a^{+\infty} xe^{-(1+y^2)x^2}\,dx = \frac{-e^{-(1+y^2)x^2}}{2(1+y^2)}\bigg|_{x=a}^{x\to+\infty} = \frac{e^{-a^2(1+y^2)}}{2(1+y^2)},$$

se deduce que

$$\tfrac{1}{2}e^{-a^2}\int_0^{+\infty}\frac{e^{-a^2y^2}}{1+y^2}\,dy = \int_a^{+\infty} e^{-x^2}\left\{\int_0^{+\infty} xe^{-x^2y^2}\,dy\right\}dx.$$

Si se introduce el cambio de variable $t = xy$, se obtiene

$$\int_0^{+\infty} xe^{-x^2y^2}\,dy = \int_0^{+\infty} e^{-t^2}\,dt = I.$$

Se deduce que

$$\int_0^{+\infty}\frac{e^{-a^2y^2}}{1+y^2}\,dy = 2e^{a^2}I\int_a^{+\infty} e^{-x^2}\,dx.$$

Si $a \to 0$, la expresión del lado derecho converge a $2I^2$. En el lado izquierdo se puede ver que el integrando está dominado por la función integrable $(1+y^2)^{-1}$. Aplicando el teorema de convergencia dominada se tiene

$$\tfrac{1}{2}\pi = \int_0^{+\infty}\frac{dy}{1+y^2} = \lim_{a\to 0}\int_0^{+\infty}\frac{e^{-a^2y^2}}{1+y^2}\,dy = 2I^2.$$

Por lo tanto, $I^2 = \pi/4$, que da una derivación de la fórmula

$$\int_0^{+\infty} e^{-x^2}\,dx = \tfrac{1}{2}\sqrt{\pi}.$$

(d) Si se integra por partes dos veces, se obtiene la fórmula

$$\text{(33.13)}\qquad \int_a^{+\infty} e^{-xy}\,\text{sen}\,x\,dx = \frac{e^{-ay}}{1+y^2}\cos a + \frac{ye^{-ay}}{1+y^2}\,\text{sen}\,a.$$

Si $x \geq a > 0$ y $y \geq \alpha > 0$, se puede argumentar igual que en el ejemplo *(b)* para demostrar que

$$\int_\alpha^{+\infty}\frac{e^{-ay}\cos a}{1+y^2}\,dy + \int_\alpha^{+\infty}\frac{ye^{-ay}\,\text{sen}\,a}{1+y^2}\,dy$$
$$= \int_a^{+\infty}\left\{\int_\alpha^{+\infty} e^{-xy}\,\text{sen}\,x\,dy\right\}dx = \int_a^{+\infty}\frac{e^{-\alpha x}\,\text{sen}\,x}{x}\,dx.$$

Se quiere tomar el límite conforme $a \to 0$. En la última integral, esto, evidentemente se puede hacer y se obtiene $\int_0^{+\infty} (e^{-\alpha x} \operatorname{sen} x/x)\, dx$. Por el hecho de que $e^{-ay} \cos a$ está dominado por 1 para $y \geq 0$, y la integral $\int_\alpha^{+\infty}(1/(1+y^2))\, dy$ existe, se puede usar el teorema de convergencia dominada 33.10 para concluir que

$$\lim_{a \to 0} \int_\alpha^{+\infty} \frac{e^{-ay} \cos a}{1+y^2}\, dy = \int_\alpha^{+\infty} \frac{dy}{1+y^2}.$$

La segunda integral es un poco más complicada ya que el mismo tipo de cálculo muestra que

$$\left| \frac{ye^{-ay} \operatorname{sen} a}{1+y^2} \right| \leq \frac{y}{1+y^2},$$

y la función dominante no es integrable, por lo tanto, se debe hacer algo mejor. Dado que $u \leq e^u$ y $|\operatorname{sen} u| \leq u$ para $u \geq 0$, se infiere que $|e^{-ay} \operatorname{sen} a| \leq 1/y$, por lo tanto, se obtiene el cálculo más preciso

$$\left| \frac{ye^{-ay} \operatorname{sen} a}{1+y^2} \right| \leq \frac{1}{1+y^2}.$$

Se puede usar ahora el teorema de convergencia dominada para tomar el límite bajo el signo de la integral para obtener

$$\lim_{a \to 0} \int_\alpha^{+\infty} \frac{ye^{-ay} \operatorname{sen} a}{1+y^2}\, dy = 0.$$

Se ha llegado a la fórmula

$$\tfrac{1}{2}\pi - \operatorname{Arc} \tan \alpha = \int_\alpha^{+\infty} \frac{dy}{1+y^2} = \int_0^{+\infty} \frac{e^{-\alpha x} \operatorname{sen} x}{x}\, dx.$$

Se quiere tomar ahora el límite conforme $\alpha \to 0$. Esta vez no se puede usar el teorema de convergencia dominada ya que $\int_0^{+\infty} x^{-1} \operatorname{sen} x\, dx$ no es absolutamente convergente. Aun cuando la convergencia de $e^{-\alpha x}$ a 1 conforme $\alpha \to 0$ es monótona, ei hecho de que sen x tome ambos signos implica que la convergencia de todo el integrando no es monótona. Afortunadamente, ya se ha visto en el ejemplo 33.5 *(d)* que la convergencia de la integral es uniforme para $\alpha \geq 0$. De acuerdo con el teorema 33.6, la integral es continua para $\alpha \geq 0$ y por lo tanto nuevamente se obtiene la fórmula

$$\int_0^{+\infty} \frac{\operatorname{sen} x}{x}\, dx = \tfrac{1}{2}\pi. \tag{33.14}$$

Ejercicios

33.A. Demostrar que la integral $\int_0^{+\infty} x^t e^{-x}\, dx$ converge uniformemente para t en un intervalo $[0, \beta]$ pero que no converge uniformemente para $t \geq 0$.

33.B. Demostrar que la integral

$$\int_0^{+\infty} \frac{\operatorname{sen}(tx)}{x}\, dx$$

es uniformemente convergente para $t \geq 1$, pero que no es absolutamente convergente para ninguno de estos valores de t.

33.C. ¿Para qué valores de t convergen uniformemente las siguientes integrales infinitas?

(a) $\int_0^{+\infty} \frac{dx}{x^2+t^2}$, (b) $\int_0^{+\infty} \frac{dx}{x^2+t}$,

(c) $\int_0^{+\infty} e^{-x} \cos tx\, dx$, (d) $\int_0^{+\infty} x^n e^{-x^2} \cos tx\, dx$,

(e) $\int_0^{+\infty} e^{-x^2-t^2/x^2}\, dx$, (f) $\int_0^{+\infty} \frac{t}{x^2} e^{-x^2-t^2/x^2}\, dx$.

33.D. Usar la fórmula (33.14) para demostrar que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

33.E. Usar la fórmula (33.14) para demostrar que $\int_0^{+\infty} e^{-tx^2}\, dx = \frac{1}{2}\sqrt{\pi/t}$ para $t>0$. Justificar la diferenciación y demostrar que

$$\int_0^{+\infty} x^{2n} e^{-x^2}\, dx = \frac{1 \cdot 3 \cdots (2n-1)}{2^{n+1}} \sqrt{\pi}.$$

33.F. Probar la existencia de la integral $\int_0^{+\infty} (1-e^{-x^2})x^{-2}\, dx$. (Observe que el integrando se puede definir como continuo en $x=0$.) Evaluar esta integral

(a) reemplazando e^{-x^2} por e^{-tx^2} y diferenciando con respecto a t;

(b) integrando $\int_1^{+\infty} e^{-tx^2}\, dx$ con respecto a t. Justificar todos los pasos.

33.G. Defina F para $t \in \mathbf{R}$ como

$$F(t) = \int_0^{+\infty} e^{-x^2} \cos tx\, dx.$$

Diferenciar con respecto a t e integrar por partes para demostrar que $F'(t) = (-1/2)tF(t)$. Encontrar $F(t)$ y, después de un cambio de variable probar la fórmula

$$\int_0^{+\infty} e^{-cx^2} \cos tx\, dx = \tfrac{1}{2}\sqrt{\pi/c}\, e^{-t^2/4c}, \qquad c>0.$$

33.H. Defina G para $t>0$ por medio de

$$G(t) = \int_0^{+\infty} e^{-x^2-t^2/x^2}\, dx.$$

Diferenciar y cambiar variables para demostrar que $G'(t) = -2G(t)$. Después encontrar $G(t)$ y probar la fórmula

$$\int_0^{+\infty} e^{-x^2-t^2/x^2}\, dx = \tfrac{1}{2}\sqrt{\pi}\, e^{-2|t|}.$$

33.I. Usar la fórmula (33.4), fórmulas trigonométricas elementales y artificios matemáticos para demostrar que

(a) $$\frac{2}{\pi}\int_0^{+\infty}\frac{\operatorname{sen} ax}{x}\,dx = 1, \qquad a>0,$$
$$= 0, \qquad a = 0,$$
$$= -1, \qquad a<0.$$

(b) $$\frac{2}{\pi}\int_0^{+\infty}\frac{\operatorname{sen} x \cos ax}{x}\,dx = 1, \qquad |a|<1,$$
$$= \tfrac{1}{2}, \qquad |a| = 1,$$
$$= 0, \qquad |a|>1.$$

(c) $$\frac{2}{\pi}\int_0^{+\infty}\frac{\operatorname{sen} x \operatorname{sen} ax}{x}\,dx = \frac{1}{\pi}\log\frac{a+1}{1-a}, \qquad |a|<1,$$
$$= \frac{1}{\pi}\log\frac{a+1}{a-1}, \qquad |a|>1,$$

(d) $$\frac{2}{\pi}\int_0^{+\infty}\left[\frac{\operatorname{sen} x}{x}\right]^2 dx = 1.$$

33.J. Para $n \in \boldsymbol{N}$ defina f_n por medio de

$$f_n(x) = 1/x, \qquad 1 \le x \le n,$$
$$= 0, \qquad x > n.$$

Cada f_n tiene una integral para $x \ge 1$ y la sucesión (f_n) es acotada, monótonamente creciente converge uniformemente a una función continua que no es integrable sobre $\{x \in \boldsymbol{R} : x \ge 1\}$.

33.K. Defina g_n por medio de

$$g_n(x) = 1/n, \qquad 0 \le x \le n^2,$$
$$= 0, \qquad x > n^2.$$

Cada g_n tiene una integral sobre $x \ge 0$ y la sucesión (g_n) es acotada y converge a una función g que tiene una integral sobre $x \ge 0$, pero, *no* es verdad que

$$\lim \int_0^{+\infty} g_n = \int_0^{+\infty} g.$$

¿Es monótona la convergencia?

33.L. Si $f(x, t) = (x-t)/(x+t)^3$, demostrar que

$$\int_1^A \left\{\int_1^{+\infty} f(x,t)\,dx\right\} dt > 0 \quad \text{para cada } A \ge 1;$$
$$\int_1^B \left\{\int_1^{+\infty} f(x,t)\,dt\right\} dx < 0 \quad \text{para cada } B \ge 1.$$

Por lo tanto, demostrar que

$$\int_1^{+\infty}\left\{\int_1^{+\infty} f(x,t)\,dx\right\} dt \ne \int_1^{+\infty}\left\{\int_1^{+\infty} f(x,t)\,dt\right\} dx.$$

33.M. Usando un argumento semejante al del ejemplo 33.15*(c)* y las fórmulas de los ejercicios 33.G y 33.H, demostrar que

$$\int_0^{+\infty} \frac{\cos ty}{1+y^2}\, dy = \frac{\pi}{2} e^{-|t|}.$$

33.N. Considerando las integrales iteradas de $e^{-(a+y)x}$ sen y sobre el cuadrante $x \geq 0$, $y \geq 0$, probar la fórmula

$$\int_0^{+\infty} \frac{e^{-ax}}{1+x^2}\, dx = \int_0^{+\infty} \frac{\operatorname{sen} y}{a+y}\, dy, \qquad a > 0.$$

Proyectos

33.α. Este proyecto trata a la **función gama,** que se introduce en el ejemplo 32.10 *(e)*. Recuerde que Γ está definida para x en $P = \{x \in \mathbf{R} : x > 0\}$ por medio de la integral

$$\Gamma(x) = \int_{0+}^{+\infty} e^{-t} t^{x-1}\, dt.$$

Ya se ha visto que esta integral converge para $x \in P$ y que $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

(a) Demostrar que Γ es continua en P.

(b) Demostrar que $\Gamma(x+1) = x\Gamma(x)$ para $x \in P$. *(Sugerencia:* integrar por partes en el intervalo $[\varepsilon, c]$.)

(c) Demostrar que $\Gamma(n+1) = n!$ para $n \in \mathbf{N}$.

(d) Demostrar que $\lim_{x \to 0+} x\Gamma(x) = 1$. Por lo tanto, se deduce que Γ no está acotada a la derecha de $x = 0$.

(e) Demostrar que Γ es diferenciable en P y que la segunda derivada siempre es positiva. (Por lo tanto, Γ es una función convexa en P.)

(f) Cambiando la variable t, demostrar que

$$\Gamma(x) = 2\int_{0+}^{+\infty} e^{-s^2} s^{2x-1}\, ds = u^x \int_{0+}^{+\infty} e^{-us} s^{x-1}\, ds.$$

33.B. Se introduce la **función beta** de Euler Defina $B(x, y)$ para x, y en $P = \{x \in \mathbf{R} : x > 0\}$ por medio de

$$B(x, y) = \int_{0+}^{1-} t^{x-1}(1-t)^{y-1}\, dt.$$

Si $x \geq 1$ y $y \geq 1$, esta integral es propia, pero si $0 < x < 1$ o $0 < y < 1$, la integral es impropia.

(a) Probar la convergencia de la integral para x, y en P.

(b) Demostrar que $B(x, y) = B(y, x)$.

(c) Demostrar que si x, y pertenecen a P, entonces,

$$B(x, y) = 2\int_{0+}^{(\pi/2)-} (\operatorname{sen} t)^{2x-1} (\cos t)^{2y-1}\, dt$$

y

$$B(x, y) = \int_{0+}^{+\infty} \frac{u^{x-1}}{(1+u)^{x+y}}\, du.$$

(d) Integrando la función positiva

$$f(t, u) = e^{-t^2-u^2} t^{2x-1} u^{2y-1}$$

sobre $\{(t, u): t^2+u^2=R^2, t \geq 0, u \geq 0\}$ y comparando esta integral con la integral sobre cuadrados inscritos y circunscritos, derivar la importante fórmula

$$B(x, y)=\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

(e) Probar las fórmulas de integración

$$\int_0^{\pi/2}(\operatorname{sen} x)^{2n}\,dx=\frac{\sqrt{\pi}\,\Gamma(n+\frac{1}{2})}{2\Gamma(n+1)}=\frac{1\cdot 3\cdot 5\cdots(2n-1)}{2\cdot 4\cdot 6\cdots(2n)}\frac{\pi}{2},$$

$$\int_0^{\pi/2}(\operatorname{sen} x)^{2n+1}\,dx=\frac{\sqrt{\pi}\,\Gamma(n+1)}{2\Gamma(n+\frac{3}{2})}=\frac{2\cdot 4\cdot 6\cdots(2n)}{1\cdot 3\cdot 5\cdot 7\cdots(2n+1)}.$$

33.γ. Este proyecto y el siguiente presentan algunas propiedades de la transformada de Laplace† que es importante tanto para matemáticas teóricas como aplicadas. Para simplificar el análisis, se restringirá la atención a funciones continuas f definadas de $\{t \in \boldsymbol{R}: t \geq 0\}$ a $\boldsymbol{R}$. ***transformada de Laplace*** de f es la función $\hat{f}$ definida en los números reales s por medio de la fórmula

$$\hat{f}(s)=\int_0^{+\infty} e^{-st}f(t)\,dt,$$

siempre que esta integral converja. Algunas veces a $\hat{f}$ se le denota por $\mathscr{L}(f)$.

(a) Suponga que existe un número real c tal que $|f(t)| \leq e^{ct}$ para una t suficientemente grande. Entonces, la integral que define la transformada de Laplace $\hat{f}$ converge para para $s>c$. Más aún converge uniformemente para $s \geq c+\delta$ si $\delta>0$.

(b) Si f satisface la condición de ser acotada del inciso *(a)*, entonces $\hat{f}$ es continua y tiene una derivada para $s>c$ dada por medio de la fórmula

$$\hat{f}'(s)=\int_0^{+\infty} e^{-st}(-t)f(t)\,dt.$$

(De modo que la derivada del transformado de Laplace de f es es el transformado de Laplace de la función $g(t)=-tf(t)$.]

(c) Demostrar por inducción que bajo la condición de ser acotada del inciso *(a)*, $\hat{f}$ tiene derivadas de todos los grados para $s>c$ y que

$$\hat{f}^{(n)}(s)=\int_0^{+\infty} e^{-st}(-t)^n f(t)\,dt.$$

(d) Suponga que f y g son funciones continuas tales que sus transformadas de Laplace $\hat{f}$ y $\hat{g}$ convergen para $s>s_0$, y tales que si a y b son números reales entonces la funcion $af+bg$ tiene una transformada de Laplace que converge para $s>s_0$ y que es igual a $a\hat{f}+b\hat{g}$.

(e) Si $a>0$ y $g(t)=f(at)$, entonces, $\hat{g}$ converge para $s>as_0$ y

$$\hat{g}(s)=\frac{1}{a}\hat{f}(s/a).$$

Análogamente, si $h(t)=(1/a)f(t/a)$, entonces $\hat{h}$ converge para $s>s_0/a$ y

† PIERRE SIMON LAPLACE (1749-1827), hijo de un ranchero normando, llegó a ser profesor en la Escuela Militar de París y fue electo para la Academia de Ciencias. Es famoso por su trabajo en mecánica celeste y probabilidad.

i314 (1494x2410x2 pcx)

$$\hat{h}(s) = \hat{f}(as).$$

(f) Suponga que la transformada de Laplace $\hat{f}$ de f existe para $s > s_0$ y defina f, para $t < 0$ igual a 0. Si $b > 0$ y si $g(t) = f(t-b)$, entonces, $\hat{g}$ converge para $s > s_0$ y

$$\hat{g}(s) = e^{-bs}\hat{f}(s).$$

Análogamente, si $h(t) = e^{bt}f(t)$ para cualquier real b, entonces $\hat{h}$ converge para $s > s_0 + b$ y

$$\hat{h}(s) = \hat{f}(s-b).$$

33.δ. Este proyecto es continuación del anterior y hace uso de sus resultados.

(a) Probar la siguiente tabla de transformados de Laplace.

$f(t)$	$\hat{f}(s)$	Intervalo de convergencia
1	$1/s$	$s > 0,$
t^n	$n!/s^{n+1}$	$s > 0,$
e^{at}	$(s-a)^{-1}$	$s > a,$
$t^n e^{at}$	$n!/(s-a)^{n+1}$	$s > a,$
$\operatorname{sen} at$	$\dfrac{a}{s^2+a^2}$	para toda
$\cos at$	$\dfrac{s}{s^2+a^2}$	para toda
$\operatorname{senh} at$	$\dfrac{a}{s^2-a^2}$	$s > a,$
$\cosh at$	$\dfrac{s}{s^2-a^2}$	$s > a,$
$\dfrac{\operatorname{sen} t}{t}$	$\operatorname{Arc tan}(1/s)$	$s > 0.$

(b) Suponga que f y f' son continuas para $t \geq 0$, que $\hat{f}$ converge para $s > s_0$ y que $e^{-st}f(t) \to 0$ y $t \to +\infty$ para toda $s > s_0$. Entonces, el transformado de Laplace de f' existe para $s > s_0$ y

$$\widehat{f'}(s) = s\hat{f}(s) - f(0).$$

(Sugerencia: integrar por partes.)

(c) Suponga que f, f' y f'' son continuas para $t \geq 0$ y que $\hat{f}$ converge para $s > s_0$. Además suponga que $e^{-st}f(t)$ y $e^{-st}f'(t)$ se acercan a 0 conforme $t \to +\infty$ para toda $s > s_0$. Entonces, la transformada de Laplace de f'' existe para $s > s_0$ y

$$\widehat{f''}(s) = s^2\hat{f}(s) - sf(0) - f'(0).$$

(d) Cuando todo el integrando o parte de él se ve que es una transformada de Laplace, la integral se puede calcular algunas veces cambiando el orden de integración. Usar este método para calcular la integral

$$\int_0^{+\infty} \frac{\operatorname{sen} s}{s}\, ds = \tfrac{1}{2}\pi.$$

(e) Se desea resolver la ecuación diferencial

$$y'(t) + 2y(t) = 3 \operatorname{sen} t, \qquad y(0) = 1.$$

i315 (1536x1130x2 pcx)

Suponga que esta ecuación tiene una solución y tal que las transformadas de Laplace de y existen para s suficientemente grande. En este caso el transformado de y debe satisfacer la ecuación

$$s\hat{y}(s) - y(0) + 2\hat{y}(s) = 4/(s-1), \qquad s > 1,$$

de donde se deduce que

$$\hat{y}(s) = \frac{s+3}{(s+2)(s-1)}.$$

Usar fracciones parciales y la tabla del inciso *(a)* para obtener $y(t) = \frac{4}{3}e^{t} - \frac{1}{3}e^{-2t}$, que se puede probar directamente que es una solución.

(f) Encontrar la solución de la ecuación

$$y'' + y' = 0, \qquad y(0) = a, \qquad y'(0) = b,$$

usando la transformada de Laplace.

(g) Demostrar que una ecuación diferencial homogénea lineal con coeficientes constantes se puede resolver usando el transformado de Laplace y la técnica de descomposición de una función racional en fracciones parciales.

VI
SERIES INFINITAS

El propósito de este capítulo es demostrar los teoremas más importantes de la teoría de series infinitas. Aunque se incluyen algunos resultados secundarios, là atención se centra en las proposiciones básicas. Si el lector desea obtener resultados más avanzados y aplicaciones, deberá consultar tratados más extensos.

En la primera sección se dan los principales teoremas referentes a convergencia de series infinitas en $\mathbf{R}^p$. Se obtienen algunos resultados de tipo general que sirven para probar la convergencia de series y justificar ciertas manipulaciones con series.

En la sección 35 se dan algunas "pruebas" familiares para la convergencia absoluta de series. Además de garantizar la convergencia de las series a las que se aplican estas pruebas, cada una de ellas da un cálculo cuantitativo de la *rapidez* de la convergencia.

En la siguiente sección se proporcionan algunas pruebas útiles de convergencia condicional y se da un breve análisis de series dobles y de multiplicación de series.

En la sección 37 se introduce el estudio de series de *funciones* y se prueban las propiedades básicas de series de potencia. En la última sección de este capítulo se prueban algunos de los principales resultados de la teoría de series de Fourier.

Sección 34 Convergencia de series infinitas

En los textos elementales algunas veces se "define" a una serie infinita como "una expresión de la forma

(34.1) $$x_1 + x_2 + \cdots + x_n + \cdots."$$

Sin embargo, esta "definición" carece de claridad, ya que no hay ningún valor particular que se pueda asociar *a priori* a este arreglo de símbolos que su-

giere un número infinito de sumas por efectuar. A pesar de que existen otras definiciones adaptables, una serie infinita se considerará igual a una sucesión de sumas parciales.

34.1 DEFINICION. Si $X=(x_n)$ es una sucesión en $\mathbf{R}^p$, entonces la **serie infinita** (o simplemente la serie) que genera X es la sucesión $S=(s_k)$ definida por

$$\begin{aligned} s_1 &= x_1, \\ s_2 &= s_1+x_2 \quad (=x_1+x_2), \\ &\cdots\cdots\cdots \\ s_k &= s_{k-1}+x_k \quad (=x_1+x_2+\cdots+x_k), \\ &\cdots\cdots\cdots \end{aligned}$$

Si S converge se hará referencia a lim S como la **suma** de series infinitas. Los elementos x_n se llaman **términos** y los elementos s_k se denominan **sumas parciales** de esta serie infinita.

Es convencional usar la expresión (34.1) o alguno de los símbolos

$$\sum (x_n), \qquad \sum_{n=1}^{\infty} (x_n), \qquad \sum_{n=1}^{\infty} x_n$$

para denotar la serie infinita generada por la sucesión $X=(x_n)$, así como para designar lim S en el caso en que esta serie infinita sea convergente. En la práctica, el uso doble de estas notaciones no crea confusión siempre que se sobreentienda que se debe probar la convergencia de la serie.

El lector deberá cuidarse de no confundir las palabras "sucesión" y "serie". En el lenguaje no matemático estas palabras se usan indistintamente; en cambio en matemáticas no son sinónimos. De acuerdo con nuestra definición una serie infinita es una sucesión S que se obtiene de una sucesión dada X según un procedimiento especial que antes se estableció. Hay muchas otras maneras de generar nuevas sucesiones y de aumentar "sumas" a la sucesión dada X. El lector deberá consultar libros relativos a ***series divergentes, series asintóticas y sumabilidad de series*** para obtener ejemplos de dichas teorías.

Una última aclaración con respecto a la notación. A pesar de que por lo general se les pone índice a los elementos de la serie con números naturales, algunas veces es más conveniente empezar con $n=0$, con $n=5$, o con $n=k$. cuando este sea el caso, las series resultantes o sus sumas se designarán como

$$\sum_{n=0}^{\infty} x_n, \qquad \sum_{n=5}^{\infty} x_n, \qquad \sum_{n=k}^{\infty} x_n.$$

En la definición 14.2 se definió la suma y la diferencia de dos sucesiones X, Y en $\mathbf{R}^p$. De manera análoga si c es un número real y si w es un elemento de $\mathbf{R}^p$, se definen las sucesiones $cX=(cx_n)$ y $(w\cdot x_n)$ en $\mathbf{R}^p$ y $\mathbf{R}$, respectivamente. Se estudiarán ahora las series generadas por estas sucesiones.

34.2 TEOREMA *(a) Si las series* $\sum (x_n)$ *y* $\sum (y_n)$ *convergen entonces la serie* $\sum (x_n + y_n)$ *converge y las sumas se relacionan por medio de la fórmula*

$$\sum (x_n + y_n) = \sum (x_n) + \sum (y_n).$$

Un resultado análogo es válido para la serie generada por $X - Y$.

(b) Si la serie $\sum (x_n)$ *es convergente, c es un número real y w es un elemento fijo de* $\mathbf{R}^p$*; entonces las series* $\sum (cx_n)$ *y* $\sum (w \cdot x_n)$ *convergen y*

$$\sum (cx_n) = c \sum (x_n), \qquad \sum (w \cdot x_n) = w \cdot \sum (x_n).$$

DEMOSTRACION. Este resultado se deduce directamente del teorema 15.16 y la definición 34.1. Q.E.D.

Se podría pensar que si las sucesiones $X = (x_n)$ y $Y = (y_n)$ generan series convergentes, entonces la sucesión $X \cdot Y = (x_n \cdot y_n)$ también genera una serie convergente. Se puede ver que esto no siempre ocurre tomando $X = Y = ((-1)^n/\sqrt{n})$ en R.

En seguida se ofrece una condición necesaria muy sencilla para la convergencia de una serie. Sin embargo, es más que suficiente.

34.3 LEMA. *Si* $\sum (x_n)$ *converge en* $\mathbf{R}^p$*, entonces, lim* $(x_n) = 0$.

DEMOSTRACION. Por definición, la convergencia de $\sum (x_n)$ significa que $\lim (s_k)$ existe, Pero dado que $x_k = s_k - s_{k-1}$, entonces $\lim (x_k) = \lim (s_k) - \lim (s_{k-1}) = 0$. Q.E.D.

A pesar de tener un alcance limitado, el siguiente resultado es de gran importancia.

34.4 TEOREMA. *Sea* (x_n) *una sucesión de números reales positivos. Entonces* $\sum (x_n)$ *converge si y sólo si la sucesión* $S = (s_k)$ *de sumas parciales es acotada. En este caso,*

$$\sum x_n = \lim (s_k) = \sup \{s_k\}.$$

DEMOSTRACION. Dado que $x_n \geq 0$, la sucesión de sumas parciales es monótonamente creciente:

$$s_1 \leq s_2 \leq \cdots \leq s_k \leq \cdots$$

si y sólo si es acotada. Q.E.D.

Dado que el siguiente criterio de Cauchy es precisamente una reformulación del teorema 16.10, se omitirá la demostración.

34.5 CRITERIO DE CAUCHY PARA SERIES. *La serie $\sum(x_n)$ en $\boldsymbol{R}^p$ converge si y sólo si para cada número $\varepsilon > 0$ hay un número natural $M(\varepsilon)$ tal que si $m \geq n \geq M(\varepsilon)$; entonces,*

$$\|s_m - s_n\| = \|x_{n+1} + x_{n+2} + \cdots + x_m\| < \varepsilon.$$

El concepto de convergencia absoluta con frecuencia es de gran importancia en lo que respecta a series, como se probará más adelante.

34.6 DEFINICION. Sea $x = (x_n)$ una sucesión en $\boldsymbol{R}^p$. Se dice que la serie $\sum(x_n)$ es **absolutamente convergente** si la serie $\sum(\|x_n\|)$ es convergente en $\boldsymbol{R}$. Se dice que una serie es **condicionalmente convergente** si es convergente pero no absolutamente convergente.

Se hace hincapié en que para series cuyos elementos son números reales positivos no hay distinción entre la convergencia ordinaria y la convergencia absoluta. Sin embargo para otras series sí puede haber alguna diferencia.

34.7 TEOREMA. *Si una serie en $\boldsymbol{R}^p$ es absolutamente convergente entonces es convergente.*

DEMOSTRACION. Por hipótesis, la serie $\sum(\|x_n\|)$ converge. Por lo tanto, de la necesidad del criterio de Cauchy 34.5 se infiere que dada $\varepsilon > 0$ hay un número natural $M(\varepsilon)$ tal que si $m \geq n \geq M(\varepsilon)$; entonces,

$$\|x_{n+1}\| + \|x_{n+2}\| + \cdots + \|x_m\| < \varepsilon.$$

De acuerdo con la desigualdad del triángulo, el lado izquierdo de esta relación domina a

$$\|x_{n+1} + x_{n+2} + \cdots + x_m\|.$$

Se aplica la suficiencia del criterio de Cauchy para concluir que $\sum(x_n)$ debe converger. Q.E.D.

34.8 EJEMPLOS. (*a*) considere la sucesión real $X = (a^n)$, que genera la **serie geométrica**

$$a + a^2 + \cdots + a^n + \cdots \tag{34.2}$$

Una condición necesaria para la convergencia es que el lim $(a^n) = 0$, lo que se requiere que $|a| < 1$. Si $m \geq n$, entonces

$$a^{n+1} + a^{n+2} + \cdots + a^m = \frac{a^{n+1} - a^{m+1}}{1-a}, \tag{34.3}$$

que se puede verificar multiplicando ambos lados por $1-a$ y observando la reducción del lado izquierdo. Por lo tanto, las sumas parciales satisfacen

$$|s_m - s_n| = |a^{n+1} + \cdots + a^m| \leq \frac{|a^{n+1}| + |a^{m+1}|}{|1-a|}, \qquad m \geq n.$$

Si $|a| < 1$, entonces $|a^{n+1}| \to 0$ de manera que el criterio de Cauchy implica que la serie geométrica (34.2) converge si y sólo si $|a| < 1$. Dejando $n = 0$ en (34.3) y pasando al límite con respecto a m se encuentra que (34.2) converge al límite $a/(1-a)$ cuando $|a| < 1$.

(*b*) Considere la serie armónica $\sum (1/n)$, que como se sabe diverge. Dado que $\lim (1/n) = 0$, no se puede usar en el lema 34.3 para probar esta divergencia; en cambio se desarrollará un argumento más delicado que se basará en el teorema 43.4. Se habrá de probar que una subsucesión de las sumas parciales no es acotada. De hecho, si $k_1 = 2$, entonces

$$s_{k_1} = \frac{1}{1} + \frac{1}{2},$$

y si $k_2 = 2^2$, entonces

$$s_{k_2} = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} = s_{k_1} + \frac{1}{3} + \frac{1}{4} > s_{k_1} + 2\left(\frac{1}{4}\right) = 1 + \frac{2}{2}.$$

Por inducción matemática se demuestra que si $k_r = 2^r$, entonces

$$s_{k_r} > s_{k_{r-1}} + 2^{r-1}\left(\frac{1}{2^r}\right) = s_{k_{r-1}} + \frac{1}{2} \geq 1 + \frac{r}{2}.$$

Por lo tanto, la subsucesión (s_{k_n}) no es acotada y la serie armónica no converge.

(*c*) Ahora se tratarán las **p-series** $\sum (1/n^p)$ en donde $0 < p \leq 1$ y se usa la desigualdad elemental $n^p \leq n$, para $n \in \mathbf{N}$. De esto se deduce que cuando $0 < p \leq 1$, entonces

$$\frac{1}{n} \leq \frac{1}{n^p}, \qquad n \in \mathbf{N}.$$

Dado que las sumas parciales de las series armónicas no son acotadas, esta desigualdad prueba que las sumas parciales de $\sum (1/n^p)$ no están acotadas para $0 < p \leq 1$, por lo que la serie diverge para estos valores de p.

(*d*) Considere la p-serie para $p > 1$. Dado que las sumas parciales son monótonas, es suficiente demostrar que alguna sucesión permanece acotada para poder demostrar la convergencia de la serie. Si $k_1 = 2^1 - 1 = 1$, entonces $s_{k_1} = 1$. Si $k_2 = 2^2 - 1 = 3$, se tiene

$$s_{k_2} = \frac{1}{1} + \left(\frac{1}{2^p} + \frac{1}{3^p}\right) < 1 + \frac{2}{2^p} = 1 + \frac{1}{2^{p-1}},$$

y si $k_3 = 2^3 - 1$, se tiene

$$s_{k_3} = s_{k_2} + \left(\frac{1}{4^p} + \frac{1}{5^p} + \frac{1}{6^p} + \frac{1}{7^p}\right) < s_{k_2} + \frac{4}{4^p} < 1 + \frac{1}{2^{p-1}} + \frac{1}{4^{p-1}}.$$

Sea $a = 1/2^{p-1}$; dado que $p > 1$, se puede ver que $0 < a < 1$..Por inducción matématica se tiene que si $k_r = 2^r - 1$; entonces

$$0 < s_{k_r} < 1 + a + a^2 + \cdots + a^{r-1}.$$

Por lo tanto, el número $1/(1-a)$ es una cota superior para las sumas parciales de la p-serie cuando $1 < p$. Del teorema 34.4 se deduce que para tales valores de p la p-serie converge.

(e) Considere la serie $\sum(1/(n^2+n))$. Usando fracciones parciales se puede escribir

$$\frac{1}{k^2+k} = \frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}.$$

Esta expresión prueba que las sumas parciales son desenvolventes y por lo tanto

$$s_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)} = \frac{1}{1} - \frac{1}{n+1}.$$

Se infiere que la sucesión (s_n) es convergente a 1.

Reordenamiento de series

Hablando comúnmente, un reordenamiento de una serie es otra serie que se obtiene de la serie dada usando todos los términos exactamente una vez pero revolviendo el orden en que se toman los términos. Por ejemplo, la serie armónica

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \cdots + \frac{1}{n} + \cdots$$

tiene los reordenamientos

$$\frac{1}{2} + \frac{1}{1} + \frac{1}{4} + \frac{1}{3} + \cdots + \frac{1}{2n} + \frac{1}{2n-1} + \cdots,$$

$$\frac{1}{1} + \frac{1}{2} + \frac{1}{4} + \frac{1}{3} + \frac{1}{5} + \frac{1}{7} + \cdots$$

El primer reordenamiento se obtiene intercambiando el primero y segundo términos, el tercero y el cuarto y así sucesivamente. El segundo reordenamiento se obtiene de la serie armónica tomando un "término impar" dos

"términos pares" tres "términos impares" etc. Es evidente que hay una infinidad de distintos reordenamientos posibles de la serie armónica.

34.9 DEFINICION. Una serie $\sum(y_m)$ en $\mathbf{R}^p$ es un **reordenamiento** de una serie $\sum(x_n)$ si existe una biyección f de $\mathbf{N}$ sobre $\mathbf{N}$ tal que $y_m = x_{f(m)}$ para toda $m \in \mathbf{N}$.

Existe una observación importante hecha por Riemann. Si $\sum(x_n)$ es una serie en $\mathbf{R}$, condicionalmente convergente (es decir, es convergente pero no absolutamente convergente) y si c es un número real arbitrario, entonces existe un reordenamiento de $\sum(x_n)$ que converge a c. La idea de la demostración de esta afirmación es muy elemental: se toman términos positivos hasta obtener una suma parcial que exceda a c, después se toman términos negativos de la serie dada hasta obtener una suma parcial de términos menores a c, etcétera. Dado que $\lim(x_n) = 0$, no es difícil ver que un reordenamiento que converja a c se puede construir.

En las manipulaciones con series por lo general es conveniente asegurarse de que los reordenamientos no alterarán la convergencia o el valor del límite.

34.10 TEOREMA DEL REORDENAMIENTO. *Sea $\sum(x_n)$ una serie absolutamente convergente en $\mathbf{R}^p$. Entonces, cualquier reordenamiento de $\sum(x_n)$ converge absolutamente al mismo valor.*

DEMOSTRACION. Sea $x = \sum(x_n)$, sea $\sum(y_m)$ un reordenamiento de $\sum(x_n)$, sea K una cota superior para las sumas parciales de $\sum(\|x_n\|)$. Es claro que si $t_r = y_1 + \cdots + y_r$ es una suma parcial de $\sum(y_m)$, entonces

$$\|y_1\| + \cdots + \|y_r\| \le K,$$

por lo que se infiere que $\sum(y_m)$ es absolutamente convergente a algún elemento y de $\mathbf{R}^p$. Se desea probar que $x = y$. Si $\varepsilon > 0$, sea $N(\varepsilon)$ tal que si $m > n \ge N(\varepsilon)$ y $s_n = x_1 + \cdots + x_n$, entonces $\|x - s_n\| < \varepsilon$ y

$$\sum_{k=n+1}^{m} \|x_k\| < \varepsilon.$$

Elija una suma parcial t_r d $\sum(y_m)$ tal que $\|y - t_r\| < \varepsilon$ y tal que cada $x_1, x_2, \ldots, x_n$ ocurra en t_r. Una vez que se haga esto, elija $m > n$ tan grande que toda y_k que aparezca en t_r también aparezca en s_m. Por lo tanto,

$$\|x - y\| \le \|x - s_m\| + \|s_m - t_r\| + \|t_r - y\| < \varepsilon + \sum_{n+1}^{m} \|x_k\| + \varepsilon < 3\varepsilon.$$

Dado que $\varepsilon > 0$ es arbitraria, se deduce que $x = y$. Q.E.D.

Ejercicios

34.A. Sea $\sum(a_n)$ una serie dada y sea $\sum(b_n)$ una serie en la que los términos son los mismos que en $\sum(a_n)$, excepto aquellos para los cuales $a_n = 0$ se han omitido. Demostrar que $\sum(a_n)$ converge a un número A si y sólo si $\sum(b_n)$ converge a A.

34.B. Demostrar que la convergencia de una serie no se afecta al cambiar un número finito de sus términos. (Desde luego, es posible cambiar la suma)

34.C. Demostrar que al agrupar los términos de una serie convergente introduciendo paréntesis que contienen un número finito de términos no se destruye la convergencia o el valor del límite. Sin embargo, al agrupar términos en una serie divergente se puede producir convergencia.

34.D. Demostrar que si una serie convergente de números reales contiene sólo un número finito de términos negativos entonces es absolutamente convergente.

34.E. Demostrar que si una serie de números reales es condicionalmente convergente entonces la serie de términos positivos es divergente y la serie de términos negativos es divergente.

34.F. Usando fracciones parciales, demostrar que

(a) $\sum_{n=0}^{\infty} \frac{1}{(\alpha+n)(\alpha+n+1)} = \frac{1}{\alpha}$ si $\alpha > 0$,

(b) $\sum_{n=1}^{\infty} \frac{1}{n(n+1)(n+2)} = \frac{1}{4}$.

34.G. Si $\sum (a_n)$ es una serie convergente de números reales, ¿entonces $\sum (a_n^2)$ es siempre convergente? Si $a_n \geq 0$, ¿es verdad que $\sum (\sqrt{a_n})$ siempre es convergente?

34.H. Si $\sum (a_n)$ es convergente y $a_n \geq 0$, ¿es $\sum (\sqrt{a_n a_{n+1}})$ convergente?

34.I. Sea $\sum (a_n)$ una serie de números estrictamente positivos y sea b_n, $n \in \mathbf{N}$, definida como $b_n = (a_1 + a_2 + \cdots + a_n)/n$. Demostrar que $\sum (b_n)$ siempre diverge.

34.J. Sea $\sum (a_n)$ convergente y sea c_n, $n \in \mathbf{N}$, definida como

$$c_n = \frac{a_1 + 2a_2 + \cdots + na_n}{n(n+1)}.$$

Entonces, $\sum (c_n)$ converge y es igual a $\sum (a_n)$.

34.K. Sea $\sum (a_n)$ una serie de números positivos monótonamente decrecientes. Demostrar que $\sum_{n=1}^{\infty} (a_n)$ converge si y sólo si la serie

$$\sum_{n=1}^{\infty} 2^n a_{2^n}$$

converge. A este resultado a menudo se le llama **prueba de condensación de Cauchy.** (*Sugerencia:* agrupar los términos en bloques como en el ejemplo 34.8(*b,d*).)

34.L. Usar la prueba de condensación de Cauchy para analizar la convergencia de la p-serie $\sum (1/n^p)$.

34.M. Usar la prueba de condensación de Cauchy para demostrar que las series

$$\sum \frac{1}{n \log n}, \qquad \sum \frac{1}{n(\log n)(\log \log n)},$$

$$\sum \frac{1}{n(\log n)(\log \log n)(\log \log \log n)}$$

son divergentes.

34.N. Demostrar que si $c > 1$, las series

$$\sum \frac{1}{n(\log n)^c}, \qquad \sum \frac{1}{n(\log n)(\log \log n)^c}$$

son convergentes.

34.O. Suponga que (a_n) es una sucesión monótonamente decreciente de números positivos. Demostrar que si la serie $\sum (a_n)$ converge, entonces $\lim (na_n) = 0$. ¿Es válido lo inverso?

34.P. Si $\lim (a_n) = 0$, entonces $\sum (a_n)$ y $\sum (a_n + 2a_{n+1})$ son ambas convergentes o ambas divergentes.

Sección 35 Pruebas para la convergencia absoluta

En la sección anterior se obtuvieron algunos resultados con respecto a las manipulaciones de series infinitas, específicamente en el caso en que las series son absolutamente convergentes. Sin embargo, exceptuando el criterio de Cauchy y el hecho de que los términos de una serie convergente convergen a cero, no se probó ninguna condición necesaria o suficiente para la convergencia de series infinitas.

Ahora se darán algunos resultados que se pueden usar para probar la convergencia o divergencia de series infinitas. Dada su importancia, se pondrá atención especial a la convergencia absoluta. Puesto que la convergencia absoluta de la serie $\sum (x_n)$ en $\boldsymbol{R}^p$ es equivalente a la convergencia de la serie $\sum (\|x_n\|)$ de elementos positivos de $\boldsymbol{R}$, es claro que los resultados que prueban la convergencia de series positivas reales son de interés particular.

La primera prueba muestra que si los términos de una serie positiva real estan regidos por los términos correspondientes de una serie convergente, entonces, la primera serie es convergente. Se obtiene una prueba de convergencia absoluta que el lector deberá formular.

35.1 PRUEBA DE COMPARACION. *Sean $X = (x_n)$ y $Y = (y_n)$ sucesiones reales positivas y suponga que para algún número natural K,*

(35.1) $$x_n \le y_n \quad \text{para } n \ge K,$$

entonces, la convergencia de $\sum (y_n)$ implica la convergencia de $\sum (x_n)$.

DEMOSTRACION. Si $m \ge n \ge \sup \{K, M(\varepsilon)\}$, entonces,

$$x_{n+1} + \cdots + x_m \le y_{n+1} + \cdots + y_m < \varepsilon,$$

de donde la afirmación es evidente. Q.E.D.

35.2 PRUEBA DE COMPARACION DEL LIMITE. *Suponga que $X = (x_n)$ y $Y = (y_n)$ son sucesiones reales positivas.*

(a) Si la relación

$$\lim (x_n/y_n) \neq 0 \tag{35.2}$$

es válida, entonces $\sum (x_n)$ *es convergente si y sólo si* $\sum (y_n)$ *es convergente.*

(b) Si el límite en (35.2) es cero y $\sum (y_n)$ *es convergente, entonces* $\sum (x_n)$ *es convergente.*

DEMOSTRACION. De (35.2) se infiere que para algún número real $c > 1$ y algún número natural K,

$$(1/c)y_n \leq x_n \leq cy_n \quad \text{para} \quad n \geq K.$$

Si se aplica dos veces la prueba de comparación 35.1 se obtiene la afirmación del inciso *(a)*. La demostración de *(b)* es análoga y se omitirá. Q.E.D.

Pruebas de la raíz y de la razón

En seguida se dará una prueba importante realizada por Cauchy.

35.3 PRUEBA DE LA RAIZ. *(a) Si* $X = (x_n)$ *es una sucesión en* $\mathbf{R}^p$ *y existen un número positivo* $r < 1$ *y un número natural* K *tales que*

$$\|x_n\|^{1/n} \leq r \quad \text{para } n \geq K, \tag{35.3}$$

entonces la serie $\sum (x_n)$ es absolutamente convergente.

(b) Si existe un número $r > 1$ *y un número natural* K *tales que*

$$\|x_n\|^{1/n} \geq r \quad \text{para } n \geq K, \tag{35.4}$$

entonces la serie $\sum (x_n)$ *es divergente.*

DEMOSTRACION. *(a)* Si (35.3) es válido, entonces se tiene $\|x_n\| \leq r^n$. Ahora, para $0 \leq r \leq 1$, la serie $\sum (r^n)$ es convergente, como se vio en el ejemplo 34.8(*a*). Por lo que de la prueba de comparación se infiere que $\sum (x_n)$ es absolutamente convergente.

(b) Si (35.4) es válido, entonces $\|x_n\| \geq r^n$. Pero, dado que $r \geq 1$, es falso que $\lim (\|x_n\|) = 0$. Q.E.D.

Además de probar la convergencia de $\sum (x_n)$, la prueba de la raíz se puede usar para obtener un cálculo de la rapidez de convergencia. Esta estimación es útil en cálculos numéricos, así como en algunos cálculos teóricos.

35.4 COROLARIO. *Si r satisface* $0 < r < 1$ *y si la sucesión* $X = (x_n)$ *satisface (35.3), entonces las sumas parciales* s_n, $n \geq K$, *se aproximan a la suma* $s = \sum (x_n)$ *de acuerdo con el cálculo.*

$$\|s - s_n\| \leq \frac{r^{n+1}}{1-r} \quad \text{para } n \geq K. \tag{35.5}$$

DEMOSTRACION. Si $m \geq n \geq K$, se tiene

$$\|s_m - s_n\| = \|x_{n+1} + \cdots + x_m\| \leq \|x_{n+1}\| + \cdots + \|x_m\| \leq r^{n+1} + \cdots + r^m < \frac{r^{n+1}}{1-r}.$$

Se toma ahora el límite con respecto a m para obtener (35.5).

A menudo es conveniente hacer uso de la siguiente variante de la prueba de la raíz.

35.5 COROLARIO. *Sea* $X = (x_n)$ *una sucesión en* $\mathbf{R}^p$ *y fíjese*

$$r = \lim (\|x_n\|^{1/n}), \tag{35.6}$$

siempre que este límite exista. Entonces, $\sum (x_n)$ *es absolutamente convergente cuando* $r < 1$ *y es divergente cuando* $r > 1$.

DEMOSTRACION. Se infiere que si el límite en (35.6) existe y es menor que 1, entonces existe un número real r_1 con $r < r_1 < 1$ y un número natural K tales que

$$\|x_n\|^{1/n} \leq r_1 \quad \text{para } n \geq K.$$

En este caso la serie es absolutamente convergente. Si este límite excede a 1, entonces hay un número natural real $r_2 > 1$ y un número natural K tales que

$$\|x_n\|^{1/n} \geq r_2 \quad \text{para } n \geq K,$$

en cuyo caso la serie es divergente. Q.E.D.

Este corolario se puede generalizar usando el límite superior en vez del límite. Los detalles se dejan como ejercicio. La siguiente prueba se debe a D'Alembert†

35.6 PRUEBA DE LA RAZON. *(a) Si* $X = (x_n)$ *es una sucesión de elementos distintos de 0 de* $\mathbf{R}^p$ *y hay un número positivo* $r < 1$ *y un número natural* K *tal es que*

$$\frac{\|x_{n+1}\|}{\|x_n\|} \leq r \quad \text{para } n \geq K, \tag{35.7}$$

Entonces la serie $\sum (x_n)$ *es absolutamente convergente.*

(b) Si existen un número $r \geq 1$ *y un número natural* K *tal que*

† JEAN LE ROND D'ALEMBERT (1717-1783) fue hijo del Caballero Destouches. Llegó a ser secretario de la Academia Francesa y el matemático líder de los enciclopedistas. Contribuyó en la dinámica y en ecuaciones diferenciales.

(35.8) $$\frac{\|x_{n+1}\|}{\|x_n\|} \geq r \quad \text{para} \quad n \geq K,$$

entonces la serie $\sum (x_n)$ *es divergente*

DEMOSTRACION. *(a)* Si (35.7) es válido, entonces un argumento elemental de inducción prueba que $\|x_{K+m}\| \leq r^m \|x_K\|$ para $m \geq 1$. Se infiere que para $n \geq K$ los términos de $\sum (x_n)$ están dominados por un múltiplo fijo de los ferminos de la serie geométrica $\sum (r^n)$ con $0 \leq r < 1$. De la prueba de comparación 35.1 se deduce que $\sum (x_n)$ es abosulutamente convergente.

(b) Si (35.8) es válido, entonces un argumento elemental de inducción prueba que $\|x_{K+m}\| \geq r^m \|x_K\|$ para $m \geq 1$. Dado que $r \geq 1$, es imposible tener $\lim (\|x_n\|) = 0$ y la serie no puede converger. Q.E.D.

35.7 COROLARIO. *Si r satisface* $0 \leq r < 1$ *y si la sucesión* $X = (x_n)$ *satisfce (35.7) para* $n \geq K$, *entonces las sumas parciales se aproximan a la suma* $s = \sum (x_n)$ *de acuerdo con el cálculo.*

(35.9) $$\|s - s_n\| \leq \frac{r}{1-r} \|x_n\| \qquad \textit{para } n \geq K.$$

DEMOSTRACION. La relación (35.7) implica que $\|x_{n+k}\| \leq r^k \|x_n\|$ cuando $n \geq K$. por lo tanto, si $m \geq n \geq K$, se tiene

$$\|s_m - s_n\| = \|x_{n+1} + \cdots + x_m\| \leq \|x_{n+1}\| + \cdots + \|x_m\|$$
$$\leq (r + r^2 + \cdots + r^{m-n}) \|x_n\| < \frac{r}{1-r} \|x_n\|.$$

De nuevo se toma el límite con respecto a m para obtener (35.9). Q.E.D.

35.8 COROLARIO. *Sea* $X = (x_n)$ *una sucesión* $\mathbf{R}^p$ *y fije*

$$r = \lim\left(\frac{\|x_{n+1}\|}{\|x_n\|}\right),$$

siempre que el límite exista. Entonces la serie $\sum (x_n)$ *es absolutamente convergente cuando* $r < 1$ *y divergente cuando* $r > 1$.

DEMOSTRACION. Suponga que el límite existe y $r < 1$. Si r_1 satisface $r < r_1 < 1$, entonces hay un número natural K tal que

$$\frac{\|x_{n+1}\|}{\|x_n\|} < r_1 \quad \text{para } n \geq K.$$

En este caso el teorema 35.6 prueba la convergencia absoluta de la serie. Si $r > 1$ y si r_2 satisface $1 < r_2 < r$, entonces hay un número natural K tal que

$$\frac{\|x_{n+1}\|}{\|x_n\|} > r_2 \qquad \text{para } n \geq K,$$

y en este caso hay divergencia. Q.E.D.

Prueba de Raabe

Si $r = 1$, fallan tanto la prueba de la razón como la prueba de la raíz y puede ocurrir que haya convergencia o divergencia. (Ver el ejemplo 35.13 *(d)* para ciertos fines es útil tener una forma más delicada de la prueba de la razón para el caso en que $r = 1$. El siguiente resultado, que se atribuye a Raabe†, por lo general es adecuado.

35.9 PRUEBA DE RAABE. *(a) Si* $X = (x_n)$ *es una sucesión de elementos distintos de cero de* $\mathbf{R}^p$ *y existen un número real* $a > 1$ *y un número natural K tales que (35.10)*

$$(35.10) \qquad \frac{\|x_{n+1}\|}{\|x_n\|} \leq 1 - \frac{a}{n} \quad \textit{para } n \geq K,$$

entonces la serie $\sum (x_n)$ *es absolutamente convergente.*

(b) Si existen un número real $a \leq 1$ *y un número natural K tales que*

$$(35.11) \qquad \frac{\|x_{n+1}\|}{\|x_n\|} \geq 1 - \frac{a}{n} \quad \textit{para } n \geq K,$$

entonces la serie $\sum (x_n)$ *no es absolutamente convergente.*

DEMOSTRACION. *(a)* Suponiendo que la relación (35.10) es válida, se tiene

$$k\,\|x_{k+1}\| \leq (k-1)\,\|x_k\| - (a-1)\,\|x_k\| \qquad \text{para } k \geq K.$$

se infiere que

$$(35.12) \qquad (k-1)\,\|x_k\| - k\,\|x_{k+1}\| \geq (a-1)\,\|x_k\| > 0 \qquad \text{para } k \geq K,$$

de donde la sucesión $(k\,\|x_{k+1}\|)$ es decreciente para $k \geq K$. Al sumar la relación (35.12) para $k = K, \ldots, n$ y observando que el lado izquierdo se reduce, se puede ver que

$$(K-1)\,\|x_K\| - n\,\|x_{n+1}\| \geq (a-1)(\|x_K\| + \cdots + \|x_n\|).$$

Esto prueba que las sumas parciales de $\sum (\|x_n\|)$ están acotadas y se demuestra la convergencia absoluta de $\sum (x_n)$.

(b) Si la relación (35.11) es válida para $n \geq K$ entonces, dado que $a \leq 1$,

† JOSEPH L. RAABE (1801-1859) nació en Ucrania y dio clases en Zurich. Hizo aportaciones tanto en geometría como en análisis.

$$n \|x_{n+1}\| \geq (n-a)\|x_n\| \geq (n-1)\|x_n\|.$$

Por lo tanto, la sucesión $(n \|x_{n+1}\|)$ es creciente para $n \geq K$ y existe un número $c > 0$ tal que

$$\|x_{n+1}\| > c/n, \qquad n \geq K.$$

Dado que diverge la serie armónica $\sum (1/n)$, entonces, $\sum (x_n)$ no puede ser absolutamente convergente. Q.E.D.

También se puede usar la prueba de Raabe para obtener información acerca de la rapidez de la convergencia.

35.10 COROLARIO. *Si $a > 1$ y si la sucesión $X = (x_n)$ satisface (35.10) entonces las sumas parciales se aproximan a la suma s de $\sum (x_k)$ de acuerdo con el cálculo.*

(35.13) $$\|s - s_n\| \leq \frac{n}{a-1}\|x_{n+1}\| \quad \text{para } n \geq K.$$

DEMOSTRACION. Sea $m > n \geq K$, sumar las desigualdades obtenidas de (35.12) para $k = n+1, \ldots, m$ para obtener

$$n \|x_{n+1}\| - m \|x_{m+1}\| \geq (a-1)(\|x_{n+1}\| + \cdots + \|x_m\|).$$

Por lo tanto, se tiene

$$\|s_m - s_n\| \leq \|x_{n+1}\| + \cdots + \|x_m\| \leq \frac{n}{a-1}\|x_{n+1}\|;$$

tomando el límite con respecto a m, se obtiene (35.13). Q.E.D.

En la aplicación de la prueba de Raabe puede ser conveniente usar la siguiente forma de límite menos rígida.

35.11 COROLARIO. *Sea $X = (x_n)$ una sucesión de elementos distintos de cero de $\mathbf{R}^p$ y fíjese*

(35.14) $$a = \lim\left(n\left(1 - \frac{\|x_{n+1}\|}{\|x_n\|}\right)\right),$$

siempre que este límite exista. Entonces $\sum (x_n)$ es absolutamente convergente cuando $a > 1$ y no es absolutamente convergente cuando $a < 1$.

DEMOSTRACION. Suponga que el límite (35.14) existe y satisface $a > 1$. si a_1 es cualquier número con $a > a_1 > 1$, entonces existe un número natural K tal que

$$a_1 < n\left(1 - \frac{\|x_{n+1}\|}{\|x_n\|}\right) \quad \text{para } n \geq K.$$

Por lo tanto, se infiere que

$$\frac{\|x_{n+1}\|}{\|x_n\|} < 1 - \frac{a_1}{n} \quad \text{para } n \geq K$$

y el teorema 35.9 asegura la convergencia absoluta de la serie. El caso en el que $a < 1$ es análogo y se omitirá. Q.E.D.

La prueba de la integral

En seguida se da una prueba muy efectiva, debida a Maclaurin†, para una serie de números positivos.

35.12 PRUEBA DE LA INTEGRAL. *Sea f una función positiva, decreciente y continua en $\{t : t \geq 1\}$. Entonces, la serie $\sum (f(n))$ converge si y sólo si la integral infinita*

$$\int_1^{+\infty} f(t)\, dt = \lim_n \left(\int_1^n f(t)\, dt \right)$$

existe. En el caso de convergencia, la suma parcial $s_n = \sum_{k=1}^n (f(k))$ y la suma $s = \sum_{k=1}^\infty (f(k))$ satisfacen la estima

$$(35.15) \qquad \int_{n+1}^{+\infty} f(t)\, dt \leq s - s_n \leq \int_n^{+\infty} f(t)\, dt.$$

DEMOSTRACION. Dado que f es positiva, continua y decreciente en el intervalo $[k-1, k]$, se infiere que

$$(35.16) \qquad f(k) \leq \int_{k-1}^{k} f(t)\, dt \leq f(k-1).$$

Sumando esta desigualdad para $k = 2, 3, \ldots, n$, se obtiene la relación

$$s_n - f(1) \leq \int_1^n f(t)\, dt \leq s_{n-1},$$

que prueba que ambos o ninguno de los límites

$$\lim (s_n), \qquad \lim \left(\int_1^n f(t)\, dt \right)$$

existe. Si existen, al sumar la relación (35.16) se obtiene para $k = n+1, \ldots, m$,

† COLIN MACLAURIN (1698-1746) fue alumno de Newton y profesor en Edimburgo. Era el matemático británico líder de su época y contribuyó tanto a la geometría como a la física matemática.

$$s_m - s_n \leq \int_n^m f(t)\, dt \leq s_{m-1} - s_{n-1},$$

de donde se infiere que

$$\int_{n+1}^{m+1} f(t)\, dt \leq s_m - s_n \leq \int_n^m f(t)\, dt.$$

Si se toma el límite con respecto a m en esta última desigualdad, se obtiene (35.15).
Q.E.D.

Se va a mostrar cómo se pueden aplicar los resultados de los teoremas 35.1 a 35.12 a las p-series que se introducen en el ejemplo 34.8(c).

35.13 EJEMPLOS. *(a)* Primero se aplicará la prueba de comparación. Sabiendo que la serie armónica $\sum (1/n)$ diverge, se puede ver que si $p \leq 1$, entonces $n^p \leq n$ y por lo tanto

$$\frac{1}{n} \leq \frac{1}{n^p}.$$

Después de usar la prueba de comparación 35.1 se concluye que la p-serie $\sum (1/n^p)$ diverge para $p \leq 1$.

(b) Considere ahora el caso $p = 2$; es decir la serie $\sum (1/n^2)$. Se compara la serie con la serie convergente $\sum [1/n(n+1)]$ del ejemplo 34.8(e). Dado que la relación

$$\frac{1}{n(n+1)} < \frac{1}{n^2}$$

es válida y los términos del lado izquierdo forman una serie convergente, no se puede aplicar directamente el teorema de comparación. Sin embargo, se podría aplicar este teorema si se comparase el n-ésimo término de $\sum [1/n(n+1)]$ con el $(n+1)$-ésimo término de $\sum (1/n^2)$ en vez de ello se aplicará la prueba de comparación de límite 35.2 y se tiene

$$\frac{1}{n(n+1)} \div \frac{1}{n^2} = \frac{n^2}{n(n+1)} = \frac{n}{n+1}.$$

Dado que el límite de este cociente es 1 y $\sum [1/n(n+1)]$ converge, entonces también la serie $\sum (1/n^2)$ converge.

(c) Considere ahora el caso $p \geq 2$. si se observa que $n^p \geq n^2$ para $p \geq 2$, entonces

$$\frac{1}{n^p} \leq \frac{1}{n^2},$$

una aplicación directa de la prueba de comparación asegura que $\sum(1/n^p)$ converge para $p \geq 2$. Alternativamente se podría aplicar la prueba de comparación del límite y tener

$$\frac{1}{n^p} \div \frac{1}{n^2} = \frac{n^2}{n^p} = \frac{1}{n^{p-2}}.$$

Si $p > 2$, esta expresión converge a 0, por lo que del corolario 35.2(b) se deduce que la serie $\sum(1/n^p)$ converge para $p \geq 2$.

Al usar la prueba de comparación no se puede obtener ninguna información con respecto a las p-series para $1 < p < 2$ a menos que sea posible encontrar una serie cuyo caracter de convergencia sea conocido y que se puede comparar con la serie de este rango.

(d) Se demuestran las pruebas de la raíz y de la razón como aplicada a las p-series. Observe que

$$\left(\frac{1}{n^p}\right)^{1/n} = (n^{-p})^{1/n} = (n^{1/n})^{-p}.$$

Se sabe que (véase el ejemplo 14.8(e)) la sucesión $(n^{1/n})$ converge a 1. Por lo tanto, se tiene

$$\lim\left(\left(\frac{1}{n^p}\right)^{1/n}\right) = 1,$$

de tal manera que la prueba de la raíz (en la forma del corolario 35.5) no es aplicable.

De la misma manera, dado que

$$\frac{1}{(n+1)^p} \div \frac{1}{n^p} = \frac{n^p}{(n+1)^p} = \frac{1}{(1+1/n)^p},$$

y puesto que la sucesión $((1+1/n)^p)$ converge a 1, la prueba de la razón (en la forma del corolario 35.8) no es aplicable.

(e) No habiendo otra alternativa, se aplica la prueba de Raabe a las p-series para valores integrales de p. Primero se intenta usar el corolario 35.11 Observe que

$$n\left(1 - \frac{(n+1)^{-p}}{n^{-p}}\right) = n\left(1 - \frac{n^p}{(n+1)^p}\right)$$
$$= n\left(1 - \frac{(n+1-1)^p}{(n+1)^p}\right) = n\left(1 - \left(1 - \frac{1}{n+1}\right)^p\right).$$

Si p es un entero, entonces se puede usar el teorema del binomio para obtener una estima para el último término. De hecho,

$$n\left(1 - \left(1 - \frac{1}{n+1}\right)^p\right) = n\left(1 - 1 + \frac{p}{n+1} - \frac{p(p-1)}{2(n+1)^2} + \cdots\right).$$

Si se toma el límite con respecto a n, se obtiene p. De modo que este corolario de la prueba de Raabe demuestra que la serie converge para valores enteros de $p \geq 2$ (y si fuese conocido el teorema del binomio para valores no enteros de p, esto se podría mejorar).

(f) Por último se aplica la prueba de la integral a las p-series. Sea $f(t) = t^{-p}$ y recuerde que

$$\int_1^n \frac{1}{t}\,dt = \log(n) - \log(1),$$

$$\int_1^n \frac{1}{t^p}\,dt = \frac{1}{1-p}(n^{1-p} - 1) \quad \text{para } p \neq 1.$$

A partir de estas relaciones se puede ver que la p-serie converge si $p > 1$ y diverge si $p \leq 1$.

Ejercicios

35.A. Probar la convergencia o la divergencia de las series cuyo n-ésimo término está dado por

(a) $\dfrac{1}{(n+1)(n+2)}$, (b) $\dfrac{n}{(n+1)(n+2)}$,

(c) $2^{-1/n}$, (d) $n/2^n$,

(e) $[n(n+1)]^{-1/2}$, (f) $[n^2(n+1)]^{-1/2}$,

(g) $n!/n^n$, (h) $(-1)^n n/(n+1)$.

35.B. Para cada una de las series del ejercicio 35.A que converja, calcular el residuo si sólo se toman cuatro términos. Si se desea determinar la suma con una aproximación menor que 1/1000, ¿cuántos términos se deben tomar?

35.C. Analizar la convergencia o la divergencia de las series cuyo n-ésimo término (para n suficientemente grande) está dado por

(a) $[\log n]^{-p}$, (b) $[\log n]^{-n}$,

(c) $[\log n]^{-\log n}$, (d) $[\log n]^{-\log\log n}$,

(e) $[n \log n]^{-1}$, (f) $[n(\log n)(\log\log n)^2]^{-1}$.

35.D. Analizar la convergencia o la divergencia de las series con n-ésimo término

(a) $2^n e^{-n}$, (b) $n^n e^{-n}$,

(c) $e^{-\log n}$, (d) $(\log n)\, e^{-\sqrt{n}}$,

(e) $n!\, e^{-n}$, (f) $n!\, e^{-n^2}$

35.E. Demostrar que la serie

$$\frac{1}{1^2} + \frac{1}{2^3} + \frac{1}{3^2} + \frac{1}{4^3} + \cdots$$

es convergente pero que ni la prueba de la raíz ni la prueba de la razón son aplicables.

35.F. Si a y b son númros positivos, entonces

$$\sum \frac{1}{(an+b)^p}$$

converge si $p > 1$ y diverge si $p \leq 1$.

35.G. Analizar las series cuyo n-ésimo término es

(a) $\dfrac{n!}{3\cdot 5\cdot 7\cdots(2n+1)}$, (b) $\dfrac{(n!)^2}{(2n)!}$,

(c) $\dfrac{2\cdot 4\cdots(2n)}{3\cdot 5\cdots(2n+1)}$, (d) $\dfrac{2\cdot 4\cdots(2n)}{5\cdot 7\cdots(2n+3)}$.

35.H. La serie dada por

$$\left(\frac{1}{2}\right)^p+\left(\frac{1\cdot 3}{2\cdot 4}\right)^p+\left(\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\right)^p+\cdots$$

converge para $p>2$ y diverge para $p\leq 2$.

35.I. Sea $X=(x_n)$ una sucesión en $\boldsymbol{R}^p$ y r dado por

$$r=\limsup(\|x_n\|^{1/n}).$$

Entonces, $\sum(x_n)$ es absolutamente convergente si $r<1$ y divergente si $r>1$. [El límite superior $u=\limsup(b_n)$ de una sucesión acotada de números reales se definió en la sección 18. Es el único número u con las propiedades (i) si $u<v$ entonces $b_n\leq v$ para toda $n\in\boldsymbol{N}$, suficientemente grande y (ii) si $w<u$, entonces $w\leq b_n$ para una infinidad de $n\in\boldsymbol{N}$.]

35.J. Sea $X=(x_n)$ una sucesión de elementos distintos de cero de $\boldsymbol{R}^p$ y sea r el número dado por $r=\limsup(\|x_{n+1}\|/\|x_n\|)$.

(a) Demostrar que si $r<1$, entonces la serie $\sum(x_n)$ es absolutamente convergente.

(b) Dar un ejemplo de una serie absolutamente convergente con $r>1$.

(c) Si $\liminf(\|x_{n+1}\|/\|x_n\|)>1$, demostrar que la serie $\sum(x_n)$ no es absolutamente convergente.

35.K. Sea $X=(x_n)$ una sucesión de elementos de $\boldsymbol{R}^p$ distintos de cero y supóngase que a está dado por $a=\limsup(n(1-\|x_{n+1}\|/\|x_n\|))$.

(a) Si $a<1$, demostrar que la serie $\sum(x_n)$ no es absolutamente convergente.

(b) Dar un ejemplo de una serie divergente con $a>1$.

(c) Si $\liminf(n(1-\|x_{n+1}\|/\|x_n\|))>1$ demostrar que la serie $\sum(x_n)$ es absolutamente convergente.

35.L. Sea $X=(x_n)$ tal que $x_n>0$ para $n\in\boldsymbol{N}$. Demostrar que la serie $\sum(x_n)$ es divergente si

$$\limsup\left((\log n)\left[n\left(1-\frac{x_{n+1}}{x_n}\right)-1\right]\right)<1.$$

35.M. Sea $x_n>0$ para $n\in\boldsymbol{N}$ y supóngase que $n(1-x_{n+1}/x_n)=a+k_n/n^p$, en donde $p>0$ y (k_n) es acotada. Entonces, la serie $\sum(x_n)$ converge si $a>1$ y diverge si $a\leq 1$.

35.N. Si $p>0$, $q>0$, entonces, la serie

$$\sum\frac{(p+1)(p+2)\cdots(p+n)}{(q+1)(q+2)\cdots(q+n)}$$

converge para $q>p+1$ y diverge para $q\leq p+1$.

35.O. Demostrar que la serie $\sum(2^n n!)^2/(2n+1)!$ es divergente.

35.P. Sean $x_n > 0$ y $r = \liminf(-\log x_n/\log n)$. Demostrar que $\sum(x_n)$ converge si $r > 1$ y diverge si $r < 1$.

35.Q. Supóngase que ninguno de los números a, b, c, es un entero negativo o cero. Demostrar que la **serie hipergeométrica**

$$\frac{ab}{1!c} + \frac{a(a+1)b(b+1)}{2!c(c+1)} + \frac{a(a+1)(a+2)b(b+1)(b+2)}{3!c(c+1)(c+2)} + \cdots$$

es absolutamente convergente para $c > a + b$ y divergente para $c \leq a + b$.

35.R. Sea $a_n > 0$ y suponga que $\sum(a_n)$ converge. Construir una serie convergente $\sum(b_n)$ con $b_n > 0$ tal que $\lim(a_n/b_n) = 0$; por lo tanto, $\sum(b_n)$ converge con menos rapidez que $\sum(a_n)$. (*Sugerencia:* sean (A_n) las sumas parciales de $\sum(a_n)$ y A su límite. Defínase $r_0 = A$, $r_n = A - A_n$ y $b_n = \sqrt{r_{n-1}} - \sqrt{r_n}$.)

35.S. Sea $a_n > 0$ y suponga que $\sum(a_n)$ diverge. Construir una serie divergente $\sum(b_n)$ con $b_n > 0$ tal que $\lim(b_n/a_n) = 0$; por lo tanto, $\sum(b_n)$ diverge con menos rapidez que $\sum(a_n)$. (*Sugerencia:* sean $b_1 = \sqrt{a_1}$ y $b_n = \sqrt{a_{n-1}} - \sqrt{a_n}$, $n > 1$.)

35.T. Sea $\{n_1, n_2, \ldots\}$ la colección de números naturales que no usan el dígito 6 en su expansión decimal. Demostrar que la serie $\sum(1/n_k)$ converge a un número menor que 90. Si $\{m_1, m_2, \ldots\}$ es la colección que termina en 6, entonces $\sum(1/m_k)$ diverge.

Proyecto

35.α. a pesar de que los productos infinitos no ocurren con tanta frecuencia como las series infinitas, son importantes en muchas investigaciones y aplicaciones. Para simplificar, se restringirá la atención a productos infinitos con términos $a_n > 0$. Si $A = (a_n)$ es una sucesión de números reales estrictamente positivos, entonces el **producto infinito**, o la **sucesión de productos parciales,** generaldos por A es la sucesión $P = (p_n)$ definida por

$$p_1 = a_1,\ p_2 = p_1 a_2 (= a_1 a_2), \ldots,$$
$$p_n = p_{n-1} a_n (= a_1 a_2 \cdots a_{n-1} a_n), \ldots.$$

Si la sucesión P es convergente a un *número distinto* de cero, entonces al lim P se le llama el **producto** del producto infinito generado por A. En este caso se dice que el producto infinito es **convergente** y se escribe

$$\prod_{n=1}^{\infty} a_n, \qquad \prod(a_n), \qquad \text{o bien} \qquad a_1 a_2 a_3 \cdots a_n \cdots$$

para denotar P y lim P.
(*Observación:* la restricción de lim $P \neq 0$ no es esencial sino convencional ya que asegura que ciertas propiedades de productos finitos se apliquen a productos infinitos.)

(*a*) Demostrar que una condición necesaria para la convergencia del producto infinito es que $\lim(a_n) = 1$.

(*b*) Demostrar que una condición necesaria y suficiente para la convergencia de

$$\prod_{n=1}^{\infty} a_n, \qquad a_n > 0, \qquad \text{es la convergencia de} \qquad \sum_{n=1}^{\infty} \log a_n.$$

(*c*) Los productos infinitos a menudo tienen términos de la forma $a_n = 1 + u_n$. Manteniendo la restricción inicial, se va a suponer que $u_n > -1$ para toda $n \in \mathbf{N}$. Si

$u_n \geq 0$, demostrar que una condición necesaria y suficiente para la convergencia del infinito es la convergencia de la serie infinita $\sum (u_n)$. (*Sugerencia:* usar la prueba de comparación del límite 35.2)

(*d*) Sea $u_n > -1$. Demostrar que si la serie infinita $\sum (u_n)$ es absolutamente convergente, entonces, el producto infinito $\prod (1+u_n)$ es convergente.

(*e*) Suponga que $u_n > -1$ y que la serie $\sum (u_n)$ es convergente. Entonces, una condición necesaria y suficiente para la convergencia del producto infinito $\prod (1+u_n)$ es la convergencia de la serie infinita $\sum (u_n^2)$. (*Sugerencia:* usar el teorema de Taylor y demostrar que existen constantes positivas A y B tales que si $|u| < \frac{1}{2}$, entonces $Au^2 \leq u - \log(1+u) \leq Bu^2$.)

Sección 36 Otros resultados para series

Las pruebas que se dan en la sección 35 tienen la propiedad de garantizar que si ciertas hipótesis se cumplen entonces la serie $\sum (x_n)$ es absolutamente convergente. Ahora se sabe que la convergencia absoluta implica convergencia ordinaria; sin embargo, analizando series especiales tales como

$$\sum \frac{(-1)^n}{n}, \qquad \sum \frac{(-1)^n}{\sqrt{n}},$$

es fácil ver que la convergencia puede ocurrir aun cuando la convergencia absoluta no se cumpla. Por lo tanto, es deseable tener una prueba que dé información acerca de la convergencia ordinaria. Existen muchas de estas pruebas que son aplicables a tipos especiales de series. Posiblemente las de mayor aplicabilidad son las debidas a Abel† y a Dirichlet.

Para demostrar estas pruebas se necesita un lema al que algunas veces se le llama **fórmula de la suma parcial,** ya que pertenece a la conocida fórmula de integración por partes. En la mayoría de las aplicaciones las sucesiones X y Y son sucesiones en $\boldsymbol{R}$, pero lo los resultados son válidos cuando X y Y son sucesiones en $\boldsymbol{R}^p$ y se usa el producto interno o cuando alguna de las dos X o Y es una sucesión real y la otra está en $\boldsymbol{R}^p$.

36.1 LEMA DE ABEL. *Sean* $X = (x_n)$ *en* $\boldsymbol{R}$ *y* $Y = (y_n)$ *en* $\boldsymbol{R}^p$ *sucesiones y desígnese a las sumas parciales de* $\sum (y_n)$ *por medio de* (s_k). *Si* $m \geq n$, *entonces*

$$(36.1) \qquad \sum_{j=n}^{m} x_j y_j = (x_{m+1}s_m - x_n s_{n-1}) + \sum_{j=n}^{m} (x_j - x_{j+1}) s_j.$$

† NIELS HENRIK ABEL (1802-1829) fue hijo de un ministro noruego pobre. Cuando tenía apenas 22 años demostró la imposibilidad de resolver la ecuación general quinta por medio de radicales. Este genio autodidacta también hizo trabajos sobresalientes en series y funciones elípticas antes de su temprana muerte por tuberculosis.

DEMOSTRACION. Se puede dar una demostración de este resultado observando que $y_j = s_j - s_{j-1}$ y aparejando los términos en cada lado de la igualdad. Los detalles se le dejan al lector. Q.E.D.

Se aplica el lema de Abel para concluir que la serie $\sum (x_n y_n)$ es convergente en un caso en que ambas series $\sum (x_n)$ y $\sum (y_n)$ pueden ser divergentes.

36.2 PRUEBA DE DIRICHLET. *Suponga que las sumas parciales de $\sum (y_n)$ están acotadas. (a) Si la sucesión $X = (x_n)$ converge a cero, y si*

$$\sum |x_n - x_{n+1}| \tag{36.2}$$

es convergente, entonces la serie $\sum (x_n y_n)$ es convergente.

(b) En particular, si $X = (x_n)$ es una sucesión decreciente de números reales positivos que converge a cero, entonces la serie $\sum (x_n y_n)$ es convergente.

DEMOSTRACION. *(a)* Suponga que $\|s_j\| < B$ para toda j. Usando (36.1), se tiene la estima

$$\left\| \sum_{j=n}^{m} x_j y_j \right\| \leq \{|x_{m+1}| + |x_n| + \sum_{j=n}^{m} |x_j - x_{j+1}|\} B. \tag{36.3}$$

Si $\lim (x_n) = 0$, los dos primeros términos de la derecha se pueden hacer arbitrariamente pequeños tomando a m y n suficientemente grandes. Además, si la serie (36.2) converge, entonces el criterio de Cauchy asegura que el último término de este lado se puede hacer menor que ε tomando $m \geq n \geq M(\varepsilon)$. Por lo tanto, el criterio de Cauchy implica que la serie $\sum (x_n y_n)$ es convergente.

(b) Si $x_1 \geq x_2 \geq \ldots$, entonces la serie en (36.2) es reducible y convergente. Q.E.D.

36.3 COROLARIO. *En el inciso (b), se tiene el error de estima*

$$\left\| \sum_{j=1}^{\infty} x_j y_j - \sum_{j=1}^{n} x_j y_j \right\| \leq 2 x_{n+1} B,$$

en donde B es una cota superior para las sumas parciales $\sum (y_j)$.

DEMOSTRACION. Esto se obtiene con facilidad de la relación (36.3). Q.E.D.

La siguiente prueba fortalece la hipótesis en la serie $\sum (y_n)$, pero debilita la hipótesis en $\sum (x_n)$.

36.4 PRUEBA DE ABEL. *Suponga que la serie $\sum (y_n)$ converge en (a) Si la sucesión $X = (x_n)$ en $\boldsymbol{R}$ es tal que*

$$\sum |x_n - x_{n+1}| \tag{36.2}$$

es convergente, entonces la serie $\sum (x_n y_n)$ *es convergente.*

(b) en particular, si la sucesión $X = (x_n)$ *es monótona y convergente a* x *en* $\mathbf{R}$, *entonces la serie* $\sum (x_n y_n)$ *es convergente.*

DEMOSTRACION. *(a)* Por hipótesis, las sumas parciales s_k de $\sum (y_n)$ convergen a algún elemento s en $\mathbf{R}^p$. Por lo tanto, hay una cota B para $\{\|s_k\| : k \in \mathbf{N}\}$, y dada $\varepsilon > 0$ hay $N_1(\varepsilon)$ tal que si $n \geq N_1(\varepsilon)$, entonces $\|s_n - s\| < \varepsilon$.

Ahora, la hipótesis de que (36.2) es convergente implica que si $n \in \mathbf{N}$, entonces

$$|x_n| \leq |x_1 + (x_2 - x_1) + \cdots + (x_n - x_{n-1})|$$
$$\leq |x_1| + \sum_{k=1}^{n-1} |x_k - x_{k+1}|$$

de manera que $|x_n| < A$ para alguna $A > 0$. Más aún, existe $N_2(\varepsilon)$ tal que si $m > n \geq N_2(\varepsilon)$, entonces

$$(36.4) \qquad |x_{m+1} - x_n| \leq \sum_{j=n}^{m} |x_{j+1} - x_j| < \varepsilon.$$

Ahora sea $N_3(\varepsilon) = \sup \{N_1(\varepsilon), N_2(\varepsilon)\}$ tal que si $m > n > N_3(\varepsilon)$, entonces se tiene

$$\begin{aligned} &\|x_{m+1}s_m - x_n s_{n-1}\| \\ &\quad \leq \|x_{m+1}s_m - x_{m+1}s\| + \|x_{m+1}s - x_n s\| + \|x_n s - x_n s_{n-1}\| \\ &\quad \leq |x_{m+1}|\,\|s_m - s\| + |x_{m+1} - x_n|\,\|s\| + |x_n|\,\|s - s_{n-1}\| \\ &\quad \leq A\varepsilon + \varepsilon B + A\varepsilon = (2A + B)\varepsilon. \end{aligned}$$

Por lo tanto, por el lema de Abel 36.1, si $m > n > N_3(\varepsilon)$, entonces se tiene

$$\begin{aligned} \left\|\sum_{j=n}^{m} x_j y_j\right\| &\leq (2A + B)\varepsilon + \left\|\sum_{j=n}^{m} (x_j - x_{j+1}) s_j\right\| \\ &\leq (2A + B)\varepsilon + \left(\sum_{j=n}^{m} |x_j - x_{j+1}|\right) B \\ &\leq 2(A + B)\varepsilon, \end{aligned}$$

en donde se usó (36.4) en el último paso. Dado que $\varepsilon > 0$ es arbitraria, la convergencia de $\sum (x_j y_j)$ está probada.

(b) Si la sucesión (x_n) es monótona y converge a x, entonces la serie (36.2) es reducible y converge a $x - x_1$ o bien a $x_1 - x$. Q.E.D.

Usando el mismo tipo de argumento se puede probar el siguiente error de estima.

36.5 COROLARIO. *En el inciso (b) se tiene el error de estima*

$$\left\| \sum_{j=1}^{\infty} x_j y_j - \sum_{j=1}^{n} x_j y_j \right\| \leq |x_{n+1}| \, \|s - s_n\| + 2B \, |x - x_{n+1}|.$$

Series alternantes

Hay una clase de series reales en particular importante condicionalmente convergentes, de manera específica, aquellas cuyos términos son alternativamente positivos y negativos.

36.6 DEFINICION. Una sucesión $X = (x_n)$ de números reales distintos de cero es **alternante** si los términos $(-1)^n x_n$, $n = 1, 2, \ldots$, son números reales todos positivos (o todos negativos). Si una sucesión $X = (x_n)$ es alternante, se dice que la serie $\sum (x_n)$ que genera es una **serie alternante.**

Es útil fijar $x_n = (-1)^n z_n$ y pedir que $z_n > 0$ (o $z_n < 0$) para toda $n = 1, 2, \ldots$. La convergencia de series alternantes se trata con facilidad cuando se pueda aplicar el siguiente resultado, demostrado por Leibniz.

36.7 PRUEBA DE SERIES ALTERNANTES. *Sea $Z = (z_n)$ una sucesión decreciente de números estrictamente positivos con $\lim (z_n) = 0$. Entonces, la serie alternante $\sum ((-1)^n z_n)$ es convergente. Más aún, si S es la suma de esta serie y s_n es la n-ésima suma parcial, entonces se tiene la estima*

$$|s - s_n| \leq z_{n+1} \tag{36.5}$$

para la rapidez de convergencia.

DEMOSTRACION. Esto se deduce inmediatamente de la prueba de Dirichlet 36.2 *(b)* si se toma $y_n = (-1)^n$, pero el error estimado dado en el corolario 36.3 no es tan preciso como (36.5). También se puede proceder directamente y demostrar por inducción matemática que si $m \geq n$, entonces

$$|s_m - s_n| = |z_{n+1} - z_{n+2} + \cdots + (-1)^{m-n-1} z_m| \leq |z_{n+1}|.$$

Esto da la convergencia así como la estima (36.5). Q.E.D.

36.8 EJEMPLOS. *(a)* La serie $\sum ((-1)^n/n)$, que algunas veces se llama **serie armónica alternante,** no es absolutamente convergente. Sin embargo, de la prueba de series alternantes se deduce que es convergente.

(b) En forma análoga la serie $\sum ((-1)^2/\sqrt{n})$ es convergente, pero no absolutamente convergente.

(c) Sea $x \in \mathbf{R}$ y sea $k \in \mathbf{Z}$. Entonces, dado que

$$2 \cos kx \operatorname{sen} \tfrac{1}{2}x = \operatorname{sen}(k+\tfrac{1}{2})x - \operatorname{sen}(k-\tfrac{1}{2})x,$$

se infiere que

$$2 \operatorname{sen} \tfrac{1}{2}x[\cos x + \cdots + \cos nx] = \operatorname{sen}(n+\tfrac{1}{2})x - \operatorname{sen} \tfrac{1}{2}x.$$

Por lo tanto, si x no es un múltiplo entero de 2π, entonces

(36.6) $$\cos x + \cdots + \cos nx = \frac{\operatorname{sen}(n+\frac{1}{2})x - \operatorname{sen}\frac{1}{2}x}{2 \operatorname{sen} \frac{1}{2}x}$$

Por lo tanto, si $x \notin \{2k\pi : k \in \boldsymbol{Z}\}$, entonces

$$|\cos x + \cdots + \cos nx| \leq \frac{1}{|\operatorname{sen} \frac{1}{2}x|}.$$

Se puede aplicar entonces la prueba de Dirichlet 36.2 *(b)* para concluir que la serie $\sum (1/n) \cos nx$ converge para toda $x \notin \{2k\pi : k \in \boldsymbol{Z}\}$. Se puede ver que esta serie diverge cuando $x = 2k\pi$ para alguna $k \in \boldsymbol{Z}$.

(d) Sean $x \in \boldsymbol{R}$ y $k \in \boldsymbol{Z}$. Entonces, dado que

$$2 \operatorname{sen} kx \operatorname{sen} \tfrac{1}{2}x = \cos(k-\tfrac{1}{2})x - \cos(k+\tfrac{1}{2})x,$$

se infiere que

$$2 \operatorname{sen} \tfrac{1}{2}x[\operatorname{sen} x + \cdots + \operatorname{sen} nx] = \cos \tfrac{1}{2}x - \cos(n+\tfrac{1}{2})x.$$

Por lo tanto, si x no es un múltiplo entero de 2π, entonces

$$\operatorname{sen} x + \cdots + \operatorname{sen} nx = \frac{\cos \frac{1}{2}x - \cos(n+\frac{1}{2})x}{2 \operatorname{sen} \frac{1}{2}x}.$$

Por lo tanto, si $x \notin \{2k\pi : k \in \boldsymbol{Z}\}$, entonces

$$|\operatorname{sen} x + \cdots + \operatorname{sen} nx| \leq \frac{1}{|\operatorname{sen} \frac{1}{2}x|}.$$

Igual que antes, la prueba de Dirichlet implica la convergencia de la serie $\sum (1/n) \operatorname{sen} nx$ para toda $x \notin \{2k\pi : k \in \boldsymbol{Z}\}$. Se puede ver que esta serie también converge cuando $x = 2k\pi$ para $k \in \boldsymbol{Z}$.

(e) Sea $Y = (y_n)$ una sucesión en $\boldsymbol{R}^2$ cuyos elementos son

$$y_1 = (1, 0), \qquad y_2 = (0, 1), \qquad y_3 = (-1, 0),$$
$$y_4 = (0, -1), \ldots, y_{n+4} = y_n, \ldots.$$

Es fácil ver que la serie $\sum (y_n)$ no converge, pero sus sumas parciales s_n son acotadas; de hecho, se tiene $\|s_n\| \leq \sqrt{2}$, por lo que la prueba de Dirichlet prueba que la serie $\sum (1/n) y_n$ es convergente en $\boldsymbol{R}^2$.

Series dobles

Algunas veces es necesario considerar sumas infinitas dependiendo de dos índices enteros. La teoría de dichas series dobles se lleva a cabo reduciéndolas a sucesiones dobles, de modo que todos los resultados de la sección 19 relacionados con sucesiones dobles se pueden interpretar para series dobles. Sin embargo, no se partirá de los resultados de la sección 19; en vez de ello la atención se limitará a series dobles absolutamente convergentes, ya que ese es el tipo de series dobles que ocurren con mayor frecuencia.

Suponga que para todo par (i, j) en $N \times N$ se tiene un elemento x_{ij} en R^p. Se define la (m, n)-**ésima suma parcial** s_{mn} como

$$s_{mn} = \sum_{j=1}^{n} \sum_{i=1}^{m} x_{ij}.$$

Por analogía con la definición 34.1, se dirá que la **serie doble** $\sum (x_{ij})$ **converge** a un elemento x en R^p para toda $\varepsilon > 0$ existe un número natural $M(\varepsilon)$ tal que si $n \geq M(\varepsilon)$ y $m \geq M(\varepsilon)$, entonces

$$\|x - s_{mn}\| < \varepsilon.$$

Por analogía con la definición 34.6, se dirá que la serie doble $\sum (x_{ij})$ es **absolutamente convergente** si la serie doble $\sum (\|x_{ij}\|)$ en R es convergente.

Queda como ejercicio demostrar que si una serie doble es absolutamente convergente entonces es convergente. Más aún, un una serie doble es absolutamente convergente si y sólo si el conjunto

$$(36.7) \qquad \left\{ \sum_{j=1}^{n} \sum_{i=1}^{m} \|x_{ij}\| : m, n \in N \right\}$$

es un conjunto acotado de números reales.

Se desea relacionar a las series dobles con las series iteradas, pero sólo se tratarán series absolutamente convergentes. El siguiente resultado es muy elemental pero da un criterio útil para la convergencia absoluta de las series dobles.

36.9 LEMA. *Suponga que la serie iterada $\sum_{j=1}^{\infty} \sum_{i=1}^{\infty} \|x_{ij}\|$ converge. Entonces la serie doble $\sum (x_{ij})$ es absolutamente convergente.*

DEMOSTRACION. Por hipótesis, cada serie $\sum_{i=1}^{\infty} \|x_{ij}\|$ converge a un número positivo a_j, $j \in N$. Más aún, la serie $\sum (a_j)$ converge a un número A. Es claro que A es una cota superior del conjunto (36.7). Q.E.D.

36.10 TEOREMA. *Suponga que la serie doble $\sum (x_{ij})$ converge absolutamente a x en R^p. Entonces las dos series iteradas*

$$(36.8) \qquad \sum_{j=1}^{\infty} \sum_{i=1}^{\infty} x_{ij}, \qquad \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} x_{ij}$$

también convergen a x.

DEMOSTRACION. Por hipótesis, existe un número real positivo A que es cota superior del conjunto en (36.7). Si n es fija, se observa que

$$\sum_{i=1}^{m} \|x_{in}\| \leq \sum_{j=1}^{n} \sum_{i=1}^{m} \|x_{ij}\| \leq A,$$

para cada m en N. Se deduce entonces que para dada $n \in N$, la serie $\sum_{i=1}^{\infty} (x_{in})$ es absolutamente convergente a un elemento y_n en R^p.

Si $\varepsilon > 0$, sea $M(\varepsilon)$ tal que si $m, n \geq M(\varepsilon)$; entonces

$$\|s_{mn} - x\| < \varepsilon, \tag{36.9}$$

Dada la relación

$$s_{mn} = \sum_{i=1}^{m} x_{i1} + \sum_{i=1}^{m} x_{i2} + \cdots + \sum_{i=1}^{m} x_{in},$$

se infiere que

$$\lim_{m} (s_{mn}) = \sum_{i=1}^{\infty} x_{i1} + \sum_{i=1}^{\infty} x_{i2} + \cdots + \sum_{i=1}^{\infty} x_{in}$$

$$= y_1 + y_2 + \cdots + y_n.$$

Si se pasa al límite en (36.9) con respecto a m, se obtiene la relación

$$\left\| \sum_{j=1}^{n} y_j - x \right\| \leq \varepsilon, \qquad n \geq M(\varepsilon).$$

Esto prueba que la primera suma iterada en (36.8) existe y es igual a x. Una demostración análoga es aplicable a la segunda suma iterada. Q.E.D.

Hay otro método para sumar series dobles que se habrá de considerar, expresamente, a lo largo de las diagonales $i + j = n$.

36.11 TEOREMA. *Suponga que la serie doble* $\sum (x_{ij})$ *converge absolutamente a* x *en* R^p. *Si se define*

$$t_k = \sum_{i+j=k} x_{ij} = x_{1,k-1} + x_{2,k-2} + \cdots + x_{k-1,1},$$

entonces la serie $\sum (t_k)$ *converge absolutamente a* x.

DEMOSTRACION. Sea A el supremo del conjunto en (36.7). Observe que

$$\sum_{k=2}^{n} \|t_k\| \leq \sum_{j=1}^{n} \sum_{i=1}^{n} \|x_{ij}\| \leq A.$$

Por lo tanto, la serie $\sum (t_k)$ es absolutamente convergente y que queda por demostrar que converge a x. Sean $\varepsilon > 0$ y M tales que

$$A - \varepsilon < \sum_{j=1}^{M} \sum_{i=1}^{M} \|x_{ij}\| \leq A.$$

Si $m, n \geq M$, entonces se sigue que $\|s_{mn} - s_{MM}\|$ no es mayor que la suma $\sum (\|x_{ij}\|)$ extendida sobre todos los pares (i, j) que satisfacen $M < i \leq m$ o bien $M < j \leq n$. Por lo tanto, $\|s_{mn} - s_{MM}\| < \varepsilon$, cuando $m, n \geq M$. De aquí se infiere que $\|x - s_{MM}\| \leq \varepsilon$. Un argumento análogo prueba que si $n \geq 2M$, entonces

$$\left\| \sum_{k=2}^{n} t_k - s_{MM} \right\| < \varepsilon,$$

por lo que se sigue que $x = \sum t_k$. Q.E.D.

Multiplicación de Cauchy

En el proceso de multiplicar dos series de potencia y de juntar los términos de acuerdo con las potencias surge naturalmente un nuevo método para generar una serie a partir de dos series dadas. De aquí que sea útil para la notación distinguir a los términos de la serie por medio de los índices 0, 1, 2, ...

36.12 DEFINICION. Si $\sum_{i=0}^{\infty} (y_i)$ y $\sum_{j=0}^{\infty} (z_j)$ son series infinitas en $\boldsymbol{R}^p$, su **producto de Cauchy** es la serie $\sum_{k=0}^{\infty} (x_k)$, en donde

$$x_k = y_0 \cdot z_k + y_1 \cdot z_{k-1} + \cdots + y_k \cdot z_0.$$

En este caso el punto designa el producto interno en $\boldsymbol{R}^p$. De manera análoga se puede definir el producto de Cauchy de una serie en $\boldsymbol{R}$ y una serie en $\boldsymbol{R}^p$.

Tal vez resulte un poco sorprendente que el producto de Cauchy de dos series convergentes pueda no converger. Sin embargo, se puede ver que la serie

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{n+1}}$$

es convergente, pero el n-ésimo término del producto de Cauchy de esta serie consigo misma es

$$(-1)^n \left[\frac{1}{\sqrt{1}\sqrt{n+1}} + \frac{1}{\sqrt{2}\sqrt{n}} + \cdots + \frac{1}{\sqrt{n+1}\sqrt{1}} \right].$$

Dado que hay $n+1$ términos en los paréntesis y cada término excede a $1/(n+2)$, los términos en el producto de Cauchy no convergen a cero. Por lo tanto, este producto de Cauchy no puede converger.

36.13 TEOREMA. *Si las series $\sum_{i=0}^{\infty} y_i$ y $\sum_{j=0}^{\infty} z_j$ convergen absolutamente a y, z en $\mathbf{R}^p$, entonces su producto de Cauchy converge absolutamente a* $y \cdot z$.

DEMOSTRACION. Si $i, j = 0, 1, 2, \ldots$, sea $x_{ij} = y_i \cdot z_j$. Las hipótesis implican que la serie iterada $\sum_{j=0}^{\infty} \sum_{i=0}^{\infty} \|x_{ij}\|$ converge. Por el lema 36.9, la serie doble $\sum (x_{ij})$ es absolutamente convergente a un número real x. Aplicando los teoremas 36.10 y 36.11 se deduce que ambas series

$$\sum_{j=0}^{\infty} \sum_{i=0}^{\infty} x_{ij}, \qquad \sum_{k=0}^{\infty} \sum_{i+j=k} x_{ij}$$

convergen a x. Se puede verificar con facilidad que la serie iterada converge a $y \cdot z$ y que la serie diagonal es el producto de Cauchy de $\sum (y_i)$ y $\sum (z_i)$.

Q.E.D.

En el caso en que $p = 1$, Mertens† demostró que la convergencia absoluta de una de las series es suficiente para implicar la convergencia del producto de Cauchy. Además, Cesàro demostró que la media aritmética de las sumas parciales del producto de Cauchy converge a yz. (Véanse los ejercicios 37.0,P)

Ejercicios

36.A. Considere la serie

$$1 - \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{5} - \frac{1}{6} - \frac{1}{7} + + - - \ldots,$$

en donde los signos vienen por pares. ¿Es convergente?

36.B. sea $a_n \in \mathbf{R}$ para $n \in \mathbf{N}$ y sea $p < q$. Si la serie $\sum (a_n/n^p)$ es convergente, entonces la serie $\sum (a_n/n^q)$ también es convergente.

36.C. Si p y q son números positivos, entonces

$$\sum (-1)^n \frac{(\log n)^p}{n^q}$$

es una serie convergente.

36.D. Analizar las series cuyos n-ésimos términos son

(a) $(-1)^n \dfrac{n^n}{(n+1)^{n+1}}$, (b) $\dfrac{n^n}{(n+1)^{n+1}}$,

† FRANZ (C.J) MERTENZ (1840-1927) estudió en Berlín y dió clases en Cracovia y Viena. Contribuyó principalmente a la geometría, la teoría de números y el álgebra.

(c) $(-1)^n \dfrac{(n+1)^n}{n^n}$, (d) $\dfrac{(n+1)}{n^{n+1}}$

36.E. Suponga que $\sum (a_n)$ es una serie convergente de números reales. Demostrar que $\sum (b_n)$ converge o bien dar un contraejemplo cuando b_n se defina como

(a) a_n/n, (b) $\sqrt{a_n}/n \quad (a_n \geq 0)$,

(c) $a_n \operatorname{sen} n$, (d) $\sqrt{a_n/n} \quad (a_n \geq 0)$,

(e) $n^{1/n} a_n$, (f) $a_n/(1+|a_n|)$.

36.F. Demostrar que la serie

$$1+\frac{1}{2}-\frac{1}{3}+\frac{1}{4}+\frac{1}{5}-\frac{1}{6}++-\cdots$$

es divergente.

36.G. Si se elimina la hipótesis de que (z_n) es decreciente, demostrar que la prueba de las series alternantes 36.7 puede no ser válida.

36.H. Para $n \in \boldsymbol{N}$, defínase c_n como

$$c_n=\frac{1}{1}+\frac{1}{2}+\cdots+\frac{1}{n}-\log n.$$

Demostrar que (c_n) es una sucesión decreciente de números positivos. El límite C de esta sucesión se llama **constante de Euler** y es aproximadamente igual a 0.577. Demostrar que si se toma

$$b_n=\frac{1}{1}-\frac{1}{2}+\frac{1}{3}-\cdots-\frac{1}{2n},$$

entonces la sucesión (b_n) converge a $\log 2$. (*Sugerencia:* $b_n = c_{2n} - c_n + \log 2$.)

36.I. Sea $\sum (a_{mn})$ la serie doble dada por

$$\begin{aligned} a_{mn} &= +1, && \text{si } m-n=1, \\ &= -1, && \text{si } m-n=-1, \\ &= 0, && \text{en los otros casos} \end{aligned}$$

Demostrar que ambas sumas iteradas existen pero no son iguales y que la suma doble no existe. Sin embargo, si (s_{mn}) denota a las sumas parciales, entonces $\lim (s_{nn})$ existe.

36.J. Demostrar que si las series dobles y las series iteradas de $\sum (a_{mn})$ existen, entonces son iguales. Demostrar que la existencia de la serie doble no implica la existencia de la serie iterada; de hecho, la existencia de la serie doble ni siquiera implica que $\lim_n (a_{mn}) = 0$ para cada m.

36.K. Demostrar que si $p>1$ y $q>1$, las series dobles

$$\sum \left(\frac{1}{m^p n^q}\right) \quad \text{y} \quad \sum \left(\frac{1}{(m^2+n^2)^p}\right)$$

son convergentes.

36. L. Separando $\sum (1/n^2)$ en sus partes impar y par, demostrar que

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = 4 \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{4}{3} \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}.$$

36.M. Si $|a|<1$ y $|b|<1$, demostrar que la serie $a+b+a^2+b^2+a^3+b^3+\cdots$ converge. ¿Cuál es el límite?

36.N. Si $\sum (a_n^2)$ y $\sum (b_n^2)$ son convergentes, entonces $\sum (a_n b_n)$ es absolutamente convergente y

$$\sum a_n b_n \le \{\sum a_n^2\}^{1/2} \{\sum b_n^2\}^{1/2}.$$

Además, $\sum (a_n + b_n)^2$ converge y

$$\{\sum (a_n + b_n)^2\}^{1/2} \le \{\sum a_n^2\}^{1/2} + \{\sum b_n^2\}^{1/2}.$$

36.O. Demostrar el teorema de Mertens: Si $\sum (a_n)$ converge absolutamente a A y $\sum (b_n)$ converge a B, entonces su producto de Cauchy converge a AB. (*Sugerencia:* Denótese las sumas parciales por medio de A_n, B_n, C_n, respectivamente. Demostrar que $\lim (C_{2n} - A_n B_n) = 0$ y $\lim (C_{2n+1} - A_n B_n) = 0$.)

36.P. Demostrar el teorema de Cesàro: Sea $\sum (a_n)$ tal que converge a A y $\sum (b_n)$ tal que converge a B y sea $\sum (c_n)$ su producto de Cauchy. Si (C_n) es la sucesión de sumas parciales de $\sum (c_n)$; entonces

$$\frac{1}{n}(C_1 + C_2 + \cdots + C_n) \to AB.$$

(Sugerencia: escribir $C_1 + \cdots + C_n = A_1 B_n + \cdots + A_n B_1$, separar esta suma en tres partes y aplicar el hecho de que $A_n \to A$ y $B_n \to B$.)

Sección 37 Series de funciones

Dada la frecuencia con que aparecen y su importancia, se presentará en seguida un análisis de series infinitas de funciones. Puesto que la convergencia de una serie infinita se maneja examinando la sucesión de sumas parciales, las preguntas concernientes a series de funciones se resuelven examinando preguntas correspondientes de sucesiones de funciones. Por este motivo, una parte de esta sección es simplemente una transformación de hechos ya demostrados para sucesiones de funciones a la terminología de series. Por ejemplo, éste es el caso para la parte de la sección que trata de series de funciones generales. Sin embargo, en la segunda parte de esta sección, en la que se analizan series de potencia, surgen ciertos factores nuevos debido al tipo especial de funciones que se tratan.

37.1 DEFINICION. Si (f_n) es una sucesión de funciones definida en un subconjunto D de $\boldsymbol{R}^p$ con valores en $\boldsymbol{R}^q$, la sucesión de **sumas parciales** (s_n) de la serie infinita $\sum (f_n)$ está definida para x en D por medio de

$$s_1(x) = f_1(x),$$
$$s_2(x) = s_1(x) + f_2(x) \qquad [= f_1(x) + f_2(x)],$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$
$$s_{n+1}(x) = s_n(x) + f_{n+1}(x) \qquad [= f_1(x) + \cdots + f_n(x) + f_{n+1}(x)],$$
$$\cdots\cdots\cdots\cdots\cdots\cdots$$

En caso de que la sucesión (s_n) converja en D a una función f, se dice que la serie infinita de funciones $\sum (f_n)$ converge a f en D. Con frecuencia se escribirá

$$\sum (f_n), \qquad \sum_{n=1}^{\infty} (f_n), \qquad \text{ó} \qquad \sum_{n=1}^{\infty} f_n$$

para denotar la serie o bien el límite de la función, cuando exista.

Si la serie $\sum (\|f_n(x)\|)$ converge para cada x en D, entonces se dice que $\sum (f_n)$ es **absolutamente convergente** en D. Si la sucesión (s_n) es uniformemente convergente a D a f, entonces se dice que $\sum (f_n)$ es **uniformemente convergente** en D, o que **converge a f uniformemente en D**.

Uno de los puntos principales en series uniformemente convergentes de funciones es la validez de los siguientes resultados que proporcionan condiciones que justifican el cambio de orden de la suma y de otras operaciones de límite.

37.2 TEOREMA. *Si f_n es continua en $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ para cada $n \in \mathbf{N}$ y si $\sum (f_n)$ converge a f uniformemente en D, entonces f es continua en D.*

Esta es una transformación directa del teorema 24.1 para series. El siguiente resultado es un transformación del teorema 31.2

37.3 TEOREMA. *Suponga que las funciones de valor real f_n, $n \in \mathbf{N}$, son Riemann-Stieltjes integrables con respectò a una función monótona g $J = [a, b]$. Si la serie $\sum (f_n)$ converge a f uniformemente en J, entonces f es Riemann-Stieltjes integrable con respecto a g y*

$$\int_a^b f\,dg = \sum_{n=1}^{\infty} \int_a^b f_n\,dg. \tag{37.1}$$

Se expresará ahora el teorema de convergencia monótona 31.4 para series.

37.4 TEOREMA. *Si las f_n son funciones positivas Riemann integrables en $J = [a, b]$ y si su suma $f = \sum (f_n)$ es Riemann integrable, entonces*

$$\int_a^b f = \sum_{n=1}^{\infty} \int_a^b f_n. \tag{37.2}$$

En seguida se pasa el teorema correspondiente a diferenciación. Aquí se hará la suposición de que las series obtenidas después de una diferenciación término por término de la serie dada convergen uniformemente. Este resultado es consecuencia inmediata del teorema 28.5.

37.5 TEOREMA. *Para cada $n \in \mathbf{N}$, sea f_n una función de valor real en $J = [a, b]$ que tiene una derivada f'_n en J. Suponga que la serie infinita $\sum (f_n)$ converge para al menos un punto de J y que la serie de derivadas $\sum (f'_n)$ converge uniformemente en J. Entonces, existe una función f de valor real en J tal que $\sum (f_n)$ converge uniformemente en J a f. Además, f tiene una derivada en J y*

$$f' = \sum f'_n. \tag{37.3}$$

Pruebas para la convergencia uniforme

Ya que se han establecido algunas consecuencias de series, se estudiarán ahora algunas pruebas que se pueden usar para demostrar la convergencia uniforme.

37.6 CRITERIO DE CAUCHY. *Sea (f_n) una sucesión de funciones de $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$. La serie infinita $\sum (f_n)$ es es uniformemente convergente en D si y sólo si para toda $\varepsilon > 0$ existe una $M(\varepsilon)$ tal que si $m \geq n \geq M(\varepsilon)$, entonces*

$$\|f_n + f_{n+1} + \cdots + f_m\|_D < \varepsilon. \tag{37.4}$$

La demostración de este resultado es inmediata a partir de 17.11, que es el criterio de Cauchy correspondiente para la convergencia uniforme de sucesiones.

37.7 PRUEBA M DE WEIERSTRASS. *Sea (M_n) una sucesión de números reales no negativos tal que $\|f_n\|_D \leq M_n$ para cada $n \in \mathbf{N}$. Si la serie infinita $\sum (M_n)$ es convergente, entonces $\sum (f_n)$ es uniformemente convergente en D.*

DEMOSTRACION. Si $m > n$, se tiene la relación

$$\|f_n + \cdots + f_m\|_D \leq \|f_n\|_D + \cdots + \|f_m\|_D \leq M_n + \cdots + M_m.$$

La afirmación se infiere de los criterios de Cauchy 34.5 y 37.6 y de la convergencia de $\sum (M_n)$. Q.E.D.

Los dos resultados siguientes son muy útiles para probar la convergencia uniforme cuando la convergencia no es absoluta. Sus demostraciones se ob-

tienen modificando las demostraciones de 36.2 y 36.4 y se dejan como ejercicios.

37.8 PRUEBA DE DIRICHLET. *Sea (f_n) una sucesión de funciones en $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ tal que las sumas parciales*

$$s_n = \sum_{j=1}^{n} f_j, \qquad n \in \mathbf{N},$$

sean todas acotadas en la norma D. Sea (φ_n) una sucesión decreciente de funciones de D a R que converge uniformemente en D a cero. Entonces, la serie $\sum (\varphi_n f_n)$ converge uniformemente en D.

37.9 PRUEBA DE ABEL. *Sea $\sum (f_n)$ una serie de funciones de $D \subseteq \mathbf{R}^p$ a $\mathbf{R}^q$ uniformemente convergente en D. Sea (φ_n) una sucesión monótona de funciones de valor real en D acotada en la norma D. Entonces, la serie $\sum (\varphi_n f_n)$ converge uniformemente en D.*

37.10 EJEMPLOS. *(a)* Considere la serie $\sum_{n=1}^{\infty} (x^n/n^2)$. Si $|x| \leq 1$, entonces $|x^n/n^2| \leq 1/n^2$. Dado que la serie $\sum (1/n^2)$ es convergente, de la prueba M de Weierstrass se infiere que la serie dada es uniformemente convergente en el intervalo $[-1, 1]$.

(b) La serie que se obtiene después de la diferenciación término por término de la serie en *(a)* es $\sum_{n=1}^{\infty} (x^{n-1}/n)$. La prueba M de Weierstrass no es aplicable en el intervalo $[-1, 1]$ de modo que no se puede aplicar el teorema 37.5. De hecho, es claro que esta serie de derivadas no es convergente para $x = 1$. Sin embargo, si $0 < r < 1$, entonces la serie geométrica $\sum (r^{n-1})$ converge. Dado que

$$\left|\frac{x^{n-1}}{n}\right| \leq r^{n-1}$$

para $|x| \leq r$, de la prueba M se infiere que la serie diferenciada es uniformemente convergente en el intervalo $[-r, r]$.

(c) Una aplicación directa de la prueba M (con $M_n = 1/n^2$) prueba que $\sum_{n=1}^{\infty} (1/n^2)$ sen nx es uniformemente convergente para toda x en $\mathbf{R}$.

(d) Dado que la serie armónica $\sum (1/n)$ diverge, no se puede aplicar la prueba M a

$$\sum_{n=1}^{\infty} (1/n) \operatorname{sen} nx. \tag{37.5}$$

Sin embargo, del análisis del ejemplo 36.8 *(d)* se deduce que si el intervalo $J = [a, b]$ está contenido en el intervalo abierto $(0, 2\pi)$, entonces las sumas parciales $s_n(x) = \sum_{k=1}^{n}$ sen kx son uniformemente acotadas en J. Dado que la sucesión $(1/n)$ decrece a cero, la prueba de Dirichlet 37.8 implica que la serie (37.5) es uniformemente convergente en J.

(e) Considere $\sum_{n=1}^{\infty} ((-1)^n/n)e^{-nx}$ en el intervalo $I = [0, 1]$. Dado que la norma de n-ésimo término en I es $1/n$, no se puede aplicar la prueba de Weierstrass. La prueba de Dirichlet se puede aplicar si se demuestra que las sumas parciales de $\sum ((-1)^n e^{-nx})$ son acotadas. De manera alternativa la prueva de Abel es aplicable ya que $\sum ((-1)^n/n))$ es convergente y la sucesión acotada (e^{-nx}) es monótonamente decreciente en I (pero no uniformemente convergente a cero).

Series de potencia

Se pasará ahora al análisis de series de potencia. Esta es un tipo importante de series de funciones y tiene ciertas propiedades que no son válidas para series generales de funciones.

37.11 DEFINICION. Una serie de funciones reales $\sum (f_n)$ se dice que es una **serie de potencia en torno a** $x = c$ si la función f_n es de la forma

$$f_n(x) = a_n(x-c)^n,$$

en donde a_n y c pertenecen a R y en donde $n = 0, 1, 2, \ldots$.

Con el objeto de simplificar la notación, sólo se tratará el caso en donde $c = 0$. Esto no implica pérdida de generalidad ya que la traslación $x' = x - c$ reduce a una serie de potencia en torno a c a una serie de potencia en torno a 0. De modo que siempre que se haga referencia a una serie de potencia se tratará de una serie de la forma

$$\sum_{n=0}^{\infty} a_n x^n = a_0 + a_1 x + \cdots + a_n x^n + \cdots. \tag{37.6}$$

Aun cuando las funciones que aparecen en (37.6) están definidas sobre todo R, no es de esperar que la serie (37.6) converja para toda x en R. Por ejemplo, usando la prueba de la razón 35.8 se puede demostrar que las series

$$\sum_{n=0}^{\infty} n!x^n, \qquad \sum_{n=0}^{\infty} x^n, \qquad \sum_{n=0}^{\infty} x^n/n!,$$

convergen para x en los conjuntos

$$\{0\}, \qquad \{x \in R : |x| < 1\}, \qquad R,$$

respectivamente. De tal modo que el conjunto en el que una serie de potencia converge puede ser chico, mediano o grande. Sin embargo, un subconjunto arbitrario de R no puede ser el conjunto preciso en que una serie de potencia converja, como se habrá de probar.

Si (b_n) es una sucesión acotada de números reales no negativos, entonces se define el **límite superior** de (b_n) como el ínfimo de aquellos números v tales

que $b_n \le v$ para toda $n \in N$ suficientemente grande. Este ínfimo está determinado de manera única y se designa por medio de

$$\lim\sup(b_n).$$

Algunas otras caracterizaciones y propiedades del límite superior de una sucesión se dieron en la sección 18, pero lo único que se necesita saber es (i) que si $v > \lim\sup(b_n)$, entonces $b_n \le v$ para toda $n \in N$, suficientemente grande y (ii) que si $w < \lim\sup(b_n)$, entonces $w \le b_n$ para una infinidad de $n \in N$.

37.12 DEFINICION. Sea $\sum(a_n x^n)$ una serie de potencias. Si la sucesión $(|a_n|^{1/n})$ es acotada, se fija $\rho = \lim\sup(|a_n|^{1/n})$; si esta sucesión no es acotada, se fija $\rho = +\infty$. Se define el **radio de convergencia** de $\sum(a_n x^n)$ como

$$\begin{aligned} R &= 0, && \text{si} \quad \rho = +\infty, \\ &= 1/\rho, && \text{si} \quad 0 < \rho < +\infty, \\ &= +\infty, && \text{si} \quad \rho = 0. \end{aligned}$$

El **intervalo de convergencia** es el intervalo abierto $(-R, R)$.

Se justificará en seguida el término "radio de convergencia".

37.13 TEOREMA DE CAUCHY-HADAMARD†. *Si R es el radio de convergencia de la serie de potencia $\sum(a_n x^n)$, entonces la serie es absolutamente convergente si $|x| < R$ y divergente si $|x| > R$.*

DEMOSTRACION. Sólo se tratará el caso en que $0 < R < +\infty$, dejando como ejercicio los casos $R = 0$ y $R = +\infty$. Si $0 < |x| < R$, entonces existe un número positivo $c < 1$ tal que $|x| < cR$. Por lo tanto, $\rho < c/|x|$ y se infiere que si n es suficientemente grande, entonces $|a_n|^{1/n} \le c/|x|$. Esto equivale a afirmar que

$$(37.7) \qquad |a_n x^n| \le c^n$$

para toda n suficientemente grande. Dado que $c < 1$, la convergencia absoluta de $\sum(a_n x^n)$ se deduce de la prueba de comparación 35.1.

Si $|x| > R = 1/\rho$, entonces hay una infinidad de $n \in N$ para las que se tiene $|a_n|^{1/n} > 1/|x|$. Por lo tanto, $|a_n x^n| > 1$ para una infinidad de n, de tal manera que la sucesión $(a_n x^n)$ no converge a cero. Q.E.D.

†JACQUES HADAMARD (1865-1963), decano de matemáticos franceses, durante mucho tiempo fue admitido en la Ecole Polytechnique con el grado más alto obtenido durante su primer siglo. Fue el sucesor de Henri Poincaré en la Academia de Ciencias y demostró el teorema del número primo en 1896 a pesar de que este teorema lo había conjeturado Gauss muchos años antes. Hadamard hizo otras aportaciones a la teoría de números, el análisis complejo, las ecuaciones diferenciales parciales e incluso a la psicología.

Se podrá observar que el teorema de Cauchy-Hadamard no hace ninguna afirmación acerca de si la serie de potencia converge cuando $|x| = R$. De hecho puede suceder cualquier cosa, como lo muestran los ejemplos

$$(37.8) \qquad \sum x^n, \qquad \sum \frac{1}{n} x^n, \qquad \sum \frac{1}{n^2} x^n,$$

Dado que $\lim (n^{1/n}) = 1$ (cf. 14.8(e)), cada una de estas series de potencia tiene radio de convergencia igual a 1. La primera serie de potencia no converge en ninguno de los puntos $x = -1$ y $x = +1$; la segunda serie converge en $x = -1$ pero diverge en $x = +1$, y la tercera serie de potencia converge en $x = -1$, así como en $x = +1$. (Encontrar una serie de potencia con $R = 1$ que converja en $x = +1$ pero que diverja en $x = -1$.)

Queda como ejercicio demostrar que el radio de convergencia de $\sum (a_n x^n)$ también está dado por

$$(37.9) \qquad \lim\left(\frac{|a_n|}{|a_{n+1}|}\right),$$

siempre que este límite exista. Con frecuencia es más conveniente usar (37.9) que la definición 37.12.

El argumento usado en la demostración del teorema de Cauchy-Hadamard da la convergencia uniforme de la serie de potencia en cualquier subconjunto compacto fijo en el intervalo de convergencia $(-R, R)$.

37.14 TEOREMA. *Sea R el radio de convergencia de $\sum (a_n x^n)$ y sea K un subconjunto compacto del intervalo de convergencia $(-R, R)$. Entonces la serie de potencia converge uniformemente en K.*

DEMOSTRACION. La compacidad de $K \subseteq (-R, R)$ implica que existe una constante positiva $c < 1$ tal que $|x| < cR$ para toda $x \in K$. (¿por qué?) Por el argumento en 37.13 se deduce que para n suficientemente grande la estima (37.7) es válida para toda $x \in K$. Dado que $c < 1$, la convergencia uniforme de $\sum (a_n x^n)$ en K es una consecuencia directa de la prueba M de Weierstrass con $M_n = c^n$. Q.E.D.

37.15 TEOREMA. *El límite de una serie de potencia es continuo en el intervalo de convergencia. Una serie de potencia se puede integrar término por término en cualquier intervalo compacto contenido en el intervalo de convergencia.*

DEMOSTRACION. Si $|x_0| < R$, entonces el resultado anterior asegura que $\sum (a_n x^n)$ converge uniformemente en cualquier vecindad compacta de x_0 contenida en $(-R, R)$. La continuidad en x_0 se sigue entonces del teorema 37.2 y la integración término por término se justifica por el teorema 37.3. Q.E.D.

En seguida se demuestra que una serie de potencia se puede diferenciar término por término. Al contrario de la situación para series generales, no es necesario suponer que la serie diferenciada sea uniformemente convergente. De modo que este resultado es más aceptable que el resultado correspondiente para la diferencia de series infinitas.

37.16 TEOREMA DE DIFERENCIACION. *Una serie de potencia se puede diferenciar término por término dentro del intervalo de convergencia. De hecho, si*

$$f(x) = \sum_{n=0}^{\infty} (a_n x^n), \quad \textit{entonces} \quad f'(x) = \sum_{n=1}^{\infty} (n a_n x^{n-1}).$$

Ambas series tienen el mismo radio de convergencia.

DEMOSTRACION. Dado que $\lim (n^{1/n}) = 1$, la sucesión $(|na_n|^{1/n})$ es acotada si y sólo si la sucesión $(|a_n|^{1/n})$ es acotada. Más aún, con facilidad se puede ver que

$$\lim \sup (|na_n|^{1/n}) = \lim \sup (|a_n|^{1/n}).$$

Por lo tanto, el radio de convergencia de las dos series es el mismo, de modo que la serie formalmente diferenciada es uniformemente convergente en cada subconjunto compacto del intervalo de convergencia. Se puede aplicar entonces el teorema 37.5 para concluir que la serie formalmente diferenciada converge a la derivada de la serie dada. Q.E.D.

Se debe observar que el teorema no hace ninguna afirmación acerca de los puntos extremos del intervalo de convergencia. Si una serie es convergente en un punto extremo entonces la serie diferenciada puede ser o no ser convergente en este punto. Por ejemplo, la serie $\sum_{n=0}^{\infty} (x^n/n^2)$ converge en ambos puntos terminales $x = -1$ $x = +1$. Sin embargo, la serie diferenciada

$$\sum_{n=1}^{\infty} \frac{x^{n-1}}{n} = \sum_{m=0}^{\infty} \frac{x^m}{m+1}$$

converge en $x = -1$ pero diverge en $x = +1$.

Aplicando en forma repetida el resultado anterior se concluye que si k es cualquier número natural, entonces la serie de potencia $\sum_{n=0}^{\infty} (a_n x^n)$ puede diferenciar término por término k veces para obtener

$$\sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n x^{n-k}. \tag{37.10}$$

Más aún, esta serie converge a $f^{(k)}$ absolutamente para $|x| < R$ y uniformemente en cualquier subconjunto compacto del intervalo de convergencia.

Si se substituye $x = 0$ en (37.10), se obtiene la importante fórmula

(37.11) $$f^{(k)}(0) = k!a_k.$$

37.17 TEOREMA DE UNICIDAD. *Si $\sum (a_n x^n)$ y $\sum (b_n x^n)$ convergen en algún intervalo $(-r, r)$, $r > 0$, a la misma función f, entonces*

$$a_n = b_n \quad \text{para toda } n \in \mathbf{N}.$$

DEMOSTRACION. Las observaciones anteriores prueban que $n!a_n = f^{(n)}(0) = n!b_n$ para $n \in \mathbf{N}$. Q.E.D.

Algunos resultados adicionales†

Existen varios resultados de algunas combinaciones algebraicas de series de potencia, pero los que tratan de substitución e inversión demuestran de manera más natural usando argumentos de análisis complejo. Por este motivo no se adentrará en estas cuestiones y será suficiente un resultado en este sentido. Por suerte es uno de los más útiles.

37.18 TEOREMA DE LA MULTIPLICACION. *Si f y g están dadas en el intervalo $(-r, r)$ por medio de las series de potencia*

$$f(x) = \sum_{n=0}^{\infty} a_n x^n, \qquad g(x) = \sum_{n=0}^{\infty} b_n x^n,$$

entonces su producto está dado en este intervalo por medio de la serie $\sum (c_n x^n)$, en donde los coeficientes (c_n) son

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} \quad \text{para } n = 0, 1, 2, \ldots.$$

DEMOSTRACION. En 37.13 se vio que si $|x| < r$, entonces las series que dan $f(x)$ y $g(x)$ son absolutamente convergentes. Si se aplica el teorema 36.13 se obtiene la conclusión deseada. Q.E.D.

El teorema de la multiplicación asegura que el radio de convergencia del producto es cuando menos r. Sin embargo, puede ser más grande, como es fácilmente demostrado.

Ya se ha visto que para que una función f pueda estar dada por una serie de potencia en un intervalo $(-r, r)$, $r > 0$, es necesario que todas las derivadas de f existan en este intervalo. Se podría pensar que esta condición también es suficiente; sin embargo, las cosas no son tan sencillas. Por ejemplo, la función f, dada por

(37.12) $$\begin{aligned} f(x) &= e^{-1/x^2}, & x &\neq 0, \\ &= 0, & x &= 0, \end{aligned}$$

†El resto de esta sección se puede omitir en la primera lectura.

se puede demostrar (véase el ejercicio 37.N) que posee derivadas de todos los grados y $f^{(n)}(0)=0$ para $n=0, 1, 2, \ldots$. Si f se puede dar en un intervalo $(-r, r)$ por una serie de potencia en torno a $x=0$, entonces del teorema de unicidad 37.17 se infiere que la serie se debe desvanecer de manera idéntica, contrario al hecho de que $f(x)\neq 0$ para $x\neq 0$.

No obstante, existen algunas condiciones suficientes útiles que se pueden dar, con el objeto de garantizar que f se pueda dar por una serie de potencia. Como ejemplo, observe que del teorema de Taylor 28.6 se infiere que si existe una constante $B>0$ tal que

$$|f^{(n)}(x)|\leq B \tag{37.13}$$

para toda $|x|<r$ y $n=0, 1, 2, \ldots$, entonces $\sum_{n=0}^{\infty} f^{(n)}(0)x^n/n!$ converge a $f(x)$ para $|x|<r$. Condiciones análogas (pero menos rigurosas) se pueden dar acerca de la magnitud de las derivadas que ofrecen la misma conclusión.

Como un ejemplo se dará un resultado elegante y útil debido a Serge Bernstein que se refiere a la expansión unilateral de una función en una serie de potencia.

37.19 TEOREMA DE BERNSTEIN. Sea f una función definida que posee derivadas de todos los grados en un intervalo $[0, r]$ y suponga que f y todas sus derivadas son positivas en el intervalo $[0, r]$. Si $0\leq x<r$, entonces $f(x)$ está dada por la expansión

$$f(x)=\sum_{n=0}^{\infty}\frac{f^{(n)}(0)}{n!}x^n.$$

DEMOSTRACION. Se hará uso de la forma integral del residuo en el teorema de Taylor dado por la relación (31.3). Si $0\leq x\leq r$, entonces

$$f(x)=\sum_{k=0}^{n-1}\frac{f^{(k)}(0)}{k!}x^k+R_n, \tag{37.14}$$

en donde se tiene la fórmula

$$R_n=\frac{x^{n-1}}{(n-1)!}\int_0^1 (1-s)^{n-1}f^{(n)}(sx)\,ds.$$

Puesto que todos los términos de la suma en (37.14) son positivos, se tiene

$$f(r)\geq\frac{r^{n-1}}{(n-1)!}\int_0^1 (1-s)^{n-1}f^{(n)}(sr)\,ds. \tag{37.15}$$

Dado que $f^{(n+1)}$ es positiva, $f^{(n)}$ es creciente en $[0, r]$; por lo tanto, si x está en este intervalo, entonces

$$0\leq R_n\leq\frac{x^{n-1}}{(n-1)!}\int_0^1 (1-s)^{n-1}f^{(n)}(sr)\,ds. \tag{37.16}$$

Combinando (37.15) y (37.16) se tiene $0 \leq R_n \leq (x/r)^{n-1} f(r)$. Por lo tanto, si $0 \leq x < r$, $\lim (R_n) = 0$. Q.E.D.

Ya se vio en el teorema 37.14 que una serie de potencia converge uniformemente en todo subconjunto compacto de su intervalo de convergencia. Sin embargo, no existe ninguna razón *a priori* para pensar que este resultado se pueda extender a los puntos extremos del intervalo de convergencia. No obstante, hay un teorema de Abel que afirma que si la convergencia ocurre en uno de los puntos extremos entonces la serie converge uniformemente fuera de estos puntos extremos.

Para simplificar la notación se habrá de suponer que el radio de convergencia de la serie es igual a 1. Esto es sin pérdida de generalidad y siempre se puede obtener dejando $x' = x/R$, que es simplemente un cambio de escala.

37.20 TEOREMA DE ABEL. *Suponga que la serie de potencia* $\sum_{n=0}^{\infty} (a_n x^n)$ *converge a* $f(x)$ *para* $|x| < 1$ *y que* $\sum_{n=0}^{\infty} (a_k)$ *converge a* A. *Entonces, la serie de potencia converge uniformemente en* $\mathbf{I} = [0, 1]$ *y*

$$\lim_{x \to 1-} f(x) = A. \tag{37.17}$$

DEMOSTRACION. La prueba de Abel 37.9, con $f_n(x) = a_n$ y $\varphi_n(x) = x^n$, es aplicable para dar la convergencia uniforme de $\sum (a_n x^n)$ en $\mathbf{I}$. Por lo tanto, el límite es continuo en $\mathbf{I}$; puesto que coincide con $f(x)$ para $0 \leq x < 1$, se infiere la relación del límite (37.17) Q.E.D.

Uno de los aspectos más interesantes acerca de este resultado es que sugiere un método para asociar un límite a series que puedan no ser convergentes. De modo que si $\sum_{n=1}^{\infty} (b_n)$ es una serie infinita, se puede formar la serie de potencia correspondiente $\sum (b_n x^n)$. Si las b_n no crecen con mucha repidez, esta serie de potencia converge a una función $B(x)$ para $|x| < 1$. Si $B(x) \to \beta$ conforme $x \to 1-$, se dice que la serie $\sum (b_n)$ es **Abel sumable** a β. Este tipo de suma es análogo (pero más poderoso) que el método de Cesâro de la media aritmética mencionado en la Sección 19 y tiene consecuencias complejas e interesantes. El contenido del teorema de Abel 37.20 es análogo al teorema 19.3; asegura que si una serie ya es convergente, entonces es Abel sumable al mismo límite. Sin embargo, lo inverso no es válido ya que la serie $\sum_{n=0}^{\infty} (-1)^n$ no es convergente, pero dado que

$$\frac{1}{1+x} = \sum_{n=0}^{\infty} (-1)^n x^n,$$

se deduce que $\sum (-1)^n$ es Abel sumable a $\frac{1}{2}$.

Algunas veces sucede que si se sabe que una serie es Abel sumable y si algunas otras condiciones se satisfacen, entonces se puede demostrar que la

serie es realmente convergente. A los teoremas de esta naturaleza se les llama **teorema de Tauberian** y a menudo son muy profundos y difíciles de demostrar. También son útiles porque hacen posible pasar de un tipo de convergencia débil a un tipo más fuerte siempre que se cumplan ciertas hipótesis adicionales.

El último teorema es el primer resultado de este tipo y lo demostró A. Tauber† en 1897. Proporciona un inverso parcial del teorema de Abel.

37.21 TEOREMA DE TAUBER. *Suponga que la serie de potencia* $\sum (a_n x^n)$ *converge a* $f(x)$ *para* $|x| < 1$ *y que* $\lim (na_n) = 0$. *Si* $\lim f(x) = A$ *conforme* $x \to 1-$, *entonces la serie* $\sum (a_n)$ *converge a* A.

DEMOSTRACION. Se desea calcular diferencias tales como $\sum^N (a_n) - A$. Para hacer esto se escribe

$$\text{(37.18)} \qquad \sum_{n=0}^{N} a_n - A = \left\{\sum_{n=0}^{N} a_n - f(x)\right\} + \{f(x) - A\}$$

$$= \sum_{n=0}^{N} a_n(1-x^n) - \sum_{N+1}^{\infty} a_n x^n + \{f(x) - A\}.$$

Dado que $0 \le x < 1$, se tiene $1 - x^n = (1-x)(1+x+\cdots+x^{n-1}) < n(1-x)$, de modo que se puede dominar el primer término del lado derecho por medio de la expresión $(1-x)\sum_{n=0}^{N} na_n$.

Por hipótesis, $\lim (na_n) = 0$; por lo tanto, el teorema 19.3 implica que

$$\lim\left(\frac{1}{m+1}\sum_{n=0}^{m} na_n\right) = 0.$$

Además, se tiene la relación $A = \lim f(x)$.

Ahora sea $\varepsilon > 0$ y elija un número natural N fijo que sea tan grande que

(i) $\left|\sum_{n=0}^{N} na_n\right| < (N+1)\varepsilon$;

(ii) $|a_n| < \dfrac{\varepsilon}{N+1}$ para toda $n \ge N$;

(iii) $|f(x_0) - A| < \varepsilon$ para $x_0 = 1 - \dfrac{1}{N+1}$.

Se determinará la magnitud de (37.18) para este valor de N y x_0. De (i), (ii), (iii) y el hecho de que $(1-x_0)(N+1) = 1$, se obtiene la estima

$$\left|\sum_{n=0}^{N} a_n - A\right| \le (1-x_0)(N+1)\varepsilon + \frac{\varepsilon}{N+1}\frac{x_0^{N+1}}{1-x_0} + \varepsilon < 3\varepsilon.$$

†ALFRED TAUBER (1866-aproximadamente 1947) fue profesor en Viena. Sus estudios fueron principalmente en análisis.

Puesto que esto se puede hacer para cada $\varepsilon > 0$, queda probada la convergencia de $\sum (a_n)$ a A. Q.E.D.

Ejercicios

37.A. Analizar la convergencia y la convergencia uniforme de las series $\sum (f_n)$, en donde $f_n(x)$ está dada por

(a) $(x^2+n^2)^{-1}$, (b) $(nx)^{-2}$, $x \neq 0$,
(c) $\operatorname{sen}(x/n^2)$, (d) $(x^n+1)^{-1}$, $x \geq 0$,
(e) $x^n(x^n+1)^{-1}$, $x \geq 0$, (f) $(-1)^n(n+x)^{-1}$, $x \geq 0$.

37.B. Si $\sum (a_n)$ es una serie absolutamente convergente, entonces la serie $\sum (a_n \operatorname{sen} nx)$ es absoluta y uniformemente convergente.

37.C. Sea (c_n) una sucesión decreciente de números positivos. Si $\sum (c_n \operatorname{sen} nx)$ es uniformemente convergente, entonces $\lim (nc_n) = 0$.

37.D. Dar los detalles de la demostración de la prueba de Dirichlet 37.8.

37.E. Dar los detalles de la demostración de la prueba de Abel 37.9

37.F. Analizar los casos $R = 0$, $R = +\infty$ en el teorema de Cauchy-Hadamard 37.13.

37.G. Demostrar que el radio de convergencia R de la serie de potencia $\sum (a_n x^n)$ está dado por $\lim (|a_n|/|a_{n+1}|)$ siempre que este límite exista. Dar un ejemplo de una serie de potencia en donde este límite no exista.

37.H. Determinar el radio de convergencia de la serie $\sum (a_n x^n)$, en donde a_n está dada por

(a) $1/n^n$, (b) $n^\alpha/n!$,
(c) $n^n/n!$, (d) $(\log n)^{-1}$, $n \geq 2$.
(e) $(n!)^2/(2n)!$, (f) $n^{-\sqrt{n}}$.

37.I. Si $a_n = 1$ cuando n es el cuadrado de un número natural y $a_n = 0$ en los otros casos, encontrar el radio de convergencia de $\sum (a_n x^n)$. Si $b_n = 1$ cuando $n = m!$ para $m \in \mathbf{N}$ y $b_n = 0$ en los otros casos, encontrar el radio de convergencia de $\sum (b_n x^n)$.

37.J. Demostrar con detalle que $\lim\sup (|na_n|^{1/n}) = \lim\sup (|a_n|^{1/n})$.

37.K. Si $0 < p \leq |a_n| \leq q$ para toda $n \in \mathbf{N}$, encontrar el radio de convergencia de $\sum (a_n x^n)$.

37.L. Sea $f(x) = \sum (a_n x^n)$ para $|x| < R$. Si $f(x) = f(-x)$ para toda $|x| < R$, demostrar que $a_n = 0$ para todas las n impares.

37.M. Demostrar que si f está definida para $|x| < r$ y existe una constante B tal que $|f^{(n)}(x)| \leq B$ para toda $|x| < r$ y $n \in \mathbf{N}$, entonces la expansión de la serie de Taylor

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$$

converge a $f(x)$ para $|x| < r$.

37.N. Demostrar por inducción que la función dada en la fórmula (37.12) tiene derivadas de todos los grados en todo punto y que todas estas derivadas se desvanecen en $x = 0$. De modo que esta función no está dada por su expansión de Taylor en torno a $x = 0$.

37.O. Dar un ejemplo de una función que sea igual a su expansión de serie de Taylor en torno a $x = 0$ para $x \geq 0$, pero que no sea igual a esta expansión para $x < 0$.

37.P. El argumento delineado en el ejercicio 28.M prueba que la forma de Lagrange del residuo se puede usar para justificar la expansión binomial general

$$(1+x)^m = \sum_{n=0}^{\infty} \binom{m}{n} x^n$$

cuando x está en el intervalo $0 \leq x < 1$. De manera análoga, el ejercicio 28.N hace válida esta expansión para $-1 < x \leq 0$, pero el argumento se basa en la forma de Cauchy del residuo y algo más complicado. Para obtener una demostración alternativa de este segundo caso, aplicar el teorema de Bernsteín a $g(x) = (1-x)^m$ para $0 \leq x < 1$.

37.Q. Considere la expansión binomial en los puntos extremos $x = \pm 1$. Demostrar que si $x = -1$, entonces la serie converge absolutamente para $m \geq 0$ y diverge para $m < 0$. En $x = +1$, la serie converge absolutamente para $m \geq 0$, converge condicionalmente para $-1 < m < 0$ y diverge para $m \leq -1$.

37.R. Sea $f(x) = \tan x$ para $|x| < \pi/2$. Usar el hecho de que f es impar y el teorema de Bernsteín para demostrar que en este intervalo f está dada por su expansión de serie de Taylor en torno a $x = 0$.

37.S. Usar el teorema de Abel para demostrar que si $f(x) = \sum (a_n x^n)$ para $|x| < R$, entonces

$$\int_0^R f(x)\, dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} R^{n+1},$$

siempre que la serie del lado derecho sea convergente aun cuando la serie original pueda no converger en $x = R$. Por ello se deduce que

$$\log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}, \qquad \frac{\pi}{4} = \sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.$$

37.T. Usando el teorema de Abel, demostrar que si la las series $\sum (a_n)$ y $\sum (b_n)$ convergen y si su producto de Cauchy $\sum (c_n)$ converge, entonces se tiene $\sum (c_n) = \sum (a_n) \cdot \sum (b_n)$.

37.U. Suponga que $a_n \geq 0$ y que $f(x) = \sum (a_n x^n)$ tiene radio de convergencia 1. Si $\sum (a_n)$ diverge, demostrar que $f(x) \to +\infty$ conforme $x \to 1-$. Usar este resultado para demostrar el elemental teorema de Tauberian: Si $a_n \geq 0$ y

$$A = \lim_{x \to 1-} \sum a_n x^n,$$

entonces $\sum (a_n)$ converge a A.

37.V. Sea $\sum_{n=0}^{\infty} (p_n)$ una serie divergente de números positivos tal que el radio de convergencia de $\sum (p_n x^n)$ sea 1. Demostrar el teorema de Appell†: Si $s = \lim (a_n/p_n)$, entonces el radio de convergencia de $\sum (a_n x^n)$ también es 1 y

$$\lim_{x \to 1-} \frac{\sum a_n x^n}{\sum p_n x^n} = s.$$

† PAUL APPELL (1855-1930) fue alumno de Hermite en la Sorbona. Hizo investigaciones en análisis complejo.

(*Sugerencia:* es suficiente considerar el caso $s = 0$. Asimismo, usar el hecho de que $\lim_{x \to 1-} [\sum (p_n x^n)]^{-1} = 0$.)

37.W. Aplicar el teorema de Apell con $p(x) = \sum_{n=0}^{\infty} (x^n)$ para obtener el teorema de Abel.

37.X. Si (a_n) es una sucesión de números reales y $a_0 = 0$, sea $s_n = a_1 + \cdots + a_n$ y sea $\sigma_n = (s_1 + \cdots + s_n)/n$. Demostrar el teorema de Frobenius‡: Si $s = \lim (\sigma_n)$ entonces

$$s = \lim_{x \to 1-} \sum_{n=0}^{\infty} a_n x^n.$$

OBSERVACION. En la terminología de la teoría de sumabilidad este resultado afirma que si una sucesión (a_n) es Cesàro sumable a s, entonces también es Abel sumable a s. (*Sugerencia:* aplicar el teorema de Appell a $p(x) = (1-x)^{-2} = \sum_{n=0}^{\infty} (nx^{n-1})$ y obsérvese que $\sum (n \cdot \sigma_n x^n) = p(x) \sum (a_n x^n)$.)

Proyectos

37.α. La teoría de series de potencia presentada en el texto se extiende a series de potencia complejas.

(a) Por las observaciones de la sección 13, todas las definiciones y los teoremas que tienen significado y validez para series en $\mathbf{R}^2$ también son válidos para series con elementos en $\mathbf{C}$. En particular, los resultados que se refieren a la convergencia absoluta se extienden con facilidad

(b) Examinar los resultados que se refieren a los reordenamientos y el producto de Cauchy para ver si se extienden a $\mathbf{C}$.

(c) Demostrar que las pruebas de comparación, raíz y razón se extienden a $\mathbf{C}$.

(d) Sea R el radio de convergencia de una serie de potencia compleja

$$\sum_{n=0}^{\infty} a_n z^n.$$

Demostrar que la serie converge absolutamente para $|z| < R$ y uniformemente en cualquier subconjunto compacto de $\{z \in \mathbf{C} : |z| < R\}$.

(e) Sean f y g funciones definidas para $D = \{z \in \mathbf{C} : |z| < r\}$ con valores en C que son los límites en D de dos series de potencia. Demostrar que si f y g coinciden en $D \cap \mathbf{R}$, entonces coinciden en todo D.

(f) Demostrar que dos series de potencia en $\mathbf{C}$ se pueden multiplicar juntas en su círculo de convergencia común.

37.β. En este proyecto se define la función exponencial en términos de series de potencia. Al hacer esto, se definirá para números complejos, así como reales.

(a) Defina E para $z \in \mathbf{C}$ por medio de la serie

$$E(z) = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Demostrar que la serie es absolutamente convergente para toda $z \in \mathbf{C}$ y que es uniformemente convergente en cualquier subconjunto acotado de $\mathbf{C}$.

‡GEORG FROBENIUS (1849-1917) fue profesor en Berlín. Se le conoce por su trabajo en álgebra, así como en análisis.

(b) Demostrar que E es una función continua de $\boldsymbol{C}$ a $\boldsymbol{C}$, que $E(0) = 1$ y que

$$E(z+w) = E(z)E(w)$$

para z, w en $\boldsymbol{C}$. (*Sugerencia:* el teorema binomial para $(z+w)^n$ es válido cuando $z, w \in \boldsymbol{C}$ y $n \in \boldsymbol{N}$.)

(c) Si x y y son números reales, definan E_1 y E_2 por medio de $E_1(x) = E(x)$, $E_2(y) = E(iy)$; por lo que $E(x+iy) = E_1(x)E_2(y)$. Demostrar que E_1 toma solamente valores reales pero que E_2 tiene algunos valores no reales. Definan C y S de $\boldsymbol{R}$ a $\boldsymbol{R}$ por medio de

$$C(y) = \operatorname{Re} E_2(y), \qquad S(y) = \operatorname{Im} E_2(y)$$

para $y \in \boldsymbol{R}$ y demostrar que

$$C(y_1+y_2) = C(y_1)C(y_2) - S(y_1)S(y_2),$$
$$S(y_1+y_2) = S(y_1)C(y_2) + C(y_1)S(y_2).$$

(d) Demostrar que C y S, como se definieron en *(c)*, tienen las series de expansión

$$C(y) = \sum_{n=0}^{\infty} \frac{(-1)^n y^{2n}}{(2n)!}, \qquad S(y) = \sum_{n=0}^{\infty} \frac{(-1)^n y^{2n+1}}{(2n+1)!}.$$

(e) Demostrar que $C' = -S$ y $S' = C$. Por lo tanto, $(C^2+S^2)' = 2CC' + 2SS' = 0$ lo que implica que C^2+S^2 es idénticamente igual a 1. En particular, esto implica que tanto C como S están acotadas en valor absoluto por 1.

(f) Inferir que la funnión E_2 de $\boldsymbol{R}$ a $\boldsymbol{C}$ satisface $E_2(0) = 1$, $E_2(y_1+y_2) = E_2(y_1)E_2(y_2)$. Por lo tanto, $E_2(-y) = 1/E_2(y)$ y $|E_2(y)| = 1$ para toda y en $\boldsymbol{R}$.

Sección 38 Series de Fourier

Se da ahora la definición de la serie de Fourier† de una función continua por partes con periodo 2π. Aun cuando el análisis es breve, se ofrecen los principales teoremas de convergencia en relación con las series de Fourier. Estos teoremas son de considerable importancia en el análisis y en sus aplicaciones a la física.

En adelante se habrá de suponer que $f: \boldsymbol{R} \to \boldsymbol{R}$ tiene periodo 2π; es decir que $f(x+2\pi) = f(x)$ para toda $x \in \boldsymbol{R}$. También se supondrá que f es **continua por partes,** es decir, f es continua excepto posiblemente para un número infinito de puntos $x_1, \ldots, x_r$ en cualquier intervalo de longitud 2π, en donde f tiene límites por la ezquierda y por la derecha:

† (J.-B.) Joseph Fourier (1768-1830) fue hijo de un sastre francés. Habiendo sido educado en un monasterio, lo dejó para dedicarse a actividades matemáticas y revolucionarias. Acompañó a Napoleón a Egipto en 1798 y más tarde fué nombrado prefecto del Departamento de Isere en el sur de Francia. Durante esta época trabajó en sus logros más famosos: la teoría matemática del calor. Su trabajo dejó una huella sobresaliente en la física matemática y ha ejercido una gran influencia en ambas materias hasta nuestros días.

$$f(x_j-)=\lim_{\substack{h\to 0\\ h>0}} f(x_j-h), \qquad f(x_j+)=\lim_{\substack{h\to 0\\ h>0}} f(x_j+h).$$

El conjunto de todas las funciones $f: \mathbf{R} \to \mathbf{R}$ que tienen periodo 2π y que son continuas por partes se designará como $PC(2\pi)$. Es fácil ver que este conjunto es un espacio vectorial con las operaciones:

$$(f+g)(x)=f(x)+g(x), \qquad (cf)(x)=cf(x), \qquad x\in \mathbf{R}.$$

Por la periodicidad de $f\in PC(2\pi)$ sólo es necesario analizar a f en un intervalo de longitud 2π; por ejemplo, se tiene

$$\int_{-\pi}^{\pi} f(x)\,dx=\int_{c}^{c+2\pi} f(x)\,dx$$

para cualquier $c\in\mathbf{R}$.

En el espacio $PC(2\pi)$ se pondrá atención en las dos normas

$$\|f\|_\infty=\sup\{|f(x)|: x\in[-\pi,\pi]\}, \qquad \|f\|_2=\left(\int_{-\pi}^{\pi} (f(x))^2\,dx\right)^{1/2},$$

que están bien definidas, ya que una función en $PC(2\pi)$ es acotada y Riemann integrable. Es un ejercicio elemental demostrar que si $f\in PC(2\pi)$, entonces

(38.1) $$\|f\|_2\le\sqrt{2\pi}\,\|f\|_\infty.$$

De esta desigualdad se deduce que la convergencia en la norma $\|\cdot\|_\infty$ (es decir, *convergencia uniforme)* implica la convergencia en la norma $\|\cdot\|_2$ (es decir *convergencia del cuadrado medio).* Sin embargo, lo inverso no es válido. (Véanse los ejercicios 31.H y 38.L.)

38.1 DEFINICION. Si $f\in PC(2\pi)$, entonces los **coeficientes de Fourier** de f son los números $a_0, a_1, a_2, \ldots, b_1, b_2, \ldots$ definidos por

(38.2) $$a_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\cos nt\,dt, \qquad b_n=\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\,\mathrm{sen}\,nt\,dt.$$

La **serie de Fourier** de f es la serie

(38.3) $$\tfrac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\,\mathrm{sen}\,nx).$$

Para indicar la relación de la serie de Fourier (38.3) con la función f, con frecuencia se escribe

$$f(x)\sim\tfrac{1}{2}a_0+\sum_{n=1}^{\infty}(a_n\cos nx+b_n\,\mathrm{sen}\,nx).$$

Es conveniente aclarar que por medio de esta notación no se pretende hacer pensar que la serie de Fourier converja a $f(x)$ en ningún punto específico x. De hecho, existen funciones continuas con periodo 2π cuyas series de Fourier son divergentes en una infinidad de puntos. (Véanse Burkhill y Burkhill, página 317, y Hewitt y Ross, página 300.)

38.2 EJEMPLOS. *(a)* Defínase $f_1 \in PC(2\pi)$ en $(-\pi, \pi]$ como $f_1(x) = -1$ para $-\pi < x < 0$ y $f_1(x) = +1$ para $0 \leq x \leq +\pi$. Queda como ejercicio demostrar que la serie de Fourier para f_1 esta dada por

$$\frac{4}{\pi}\left[\frac{\operatorname{sen} x}{1} + \frac{\operatorname{sen} 3x}{3} + \frac{\operatorname{sen} 5x}{5} + \cdots\right].$$

Se demostrará que esta serie de Fourier efectivamente converge a f_1 para $0 < |x| < \pi$, pero *no* converge f_1 en $x = 0$, $\pm\pi$. (¿por que?) Obsérvese que f_1 es continua por partes pero no es continua por los puntos de $\{n\pi : n \in \mathbf{Z}\}$.

(b) Defina $f_2 \in PC(2\pi)$ en $(-\pi, \pi]$ como $f_2(x) = |x|$. Queda como ejercicio demostrar que la serie de Fourier para f_2 está dada por

$$\frac{\pi}{2} - \frac{4}{\pi}\left[\frac{\cos x}{1^2} + \frac{\cos 3x}{3^2} + \frac{\cos 5x}{5^2} + \cdots\right].$$

Es claro que esta serie converge uniformemente en $\mathbf{R}$ y se demostrará mas adelante que converge a f_2.

(c) Sea $f \in PC(2\pi)$ par, es decir $f(-x) = f(x)$ para toda $x \in \mathbf{R}$. Para una función tal, los coeficientes de Fourier $b_n = 0$ para $n = 1, 2, \ldots$, mientras que

$$a_n = \frac{2}{\pi}\int_0^{\pi} f(t) \cos nt\, dt, \qquad n = 0, 1, 2, \ldots.$$

(Observe que la función en *(b)* es par.)

(d) Sea $g \in PC(2\pi)$ impar, es decir $g(-x) = -g(x)$ para toda $x \in \mathbf{R}$. Para una función tal los coeficientes de Fourier $a_n = 0$ para $n = 0, 1, 2, \ldots$, mientras que

$$b_n = \frac{2}{\pi}\int_0^{\pi} g(t) \operatorname{sen} nt\, dt, \qquad n = 1, 2, \ldots.$$

(Observe que la función en *(a)* es impar.)

(e) Sea f continua en $\mathbf{R}$ con periodo 2π y sea su derivada f' continua por partes en $\mathbf{R}$ (y con periodo 2π.) Se van a relacionar los coeficientes de Fourier a_n, b_n de f con los coeficientes de Fourier a_n', b_n' de f' para $n = 1, 2, \ldots$. De hecho, integrando por partes se tiene

$$\begin{aligned} a_n' &= \frac{1}{\pi}\int_{-\pi}^{\pi} f'(t) \cos nt\, dt \\ &= \frac{1}{\pi}\left[f(t) \cos nt \Big|_{-\pi}^{\pi} - \int_{-\pi}^{\pi} f(t)(-n) \operatorname{sen} nt\, dt\right]. \end{aligned}$$

Si se usa el hecho de que $t \mapsto f(t) \cos nt$ tiene periodo 2π se puede ver que el primer término se desvanece y entonces $a_n' = nb_n$ para $n = 1, 2, \ldots$. en forma análoga se demuestra que $b_n' = -na_n$ para $n = 1, 2, \ldots$, (Se puede ver que si f_1, f_2 son las funciones en *(a)* y *(b)*, entonces $f_1(x) = f_2'(x)$ para $x \notin \{n\pi : n \in \mathbf{Z}\}$ y que los coeficientes de Fourier para f_1 y f_2 para $n = 1, 2, \ldots$ satisfacen las relaciones anteriores.)

En el siguiente lema se calculará el cuadrado de la distancia relativa a la norma $\|\cdot\|_2$ de f en $PC(2\pi)$ a una función arbitraria T_n de la forma

$$(38.4) \qquad T_n(x) = \tfrac{1}{2}\alpha_0 + \sum_{k=1}^{n} (\alpha_k \cos kx + \beta_k \operatorname{sen} kx);$$

dicha función algunas veces se llama **polinomio trigonométrico de grado** n. Al hacer este cálculo, es útil hacer uso de las relaciones

$$\int_{-\pi}^{\pi} (\cos kx)^2\, dx = \int_{-\pi}^{\pi} (\operatorname{sen} kx)^2\, dx = \pi, \qquad k \in \mathbf{N},$$

$$\int_{-\pi}^{\pi} \operatorname{sen} kx \operatorname{sen} nx\, dx = \int_{-\pi}^{\pi} \cos kx \cos nx\, dx = 0, \qquad k, n \in \mathbf{N}, k \neq n,$$

$$\int_{-\pi}^{\pi} \operatorname{sen} kx \cos mx\, dx = 0, \qquad k, m = 0, 1, 2, \ldots.$$

38.3 LEMA. *Si $f \in PC(2\pi)$ y T_n es un polinomio trigonométrico de grado n (es decir T_n es de la forma (38.4), entonces*

$$(38.5) \quad \|f - T_n\|_2^2 = \|f\|_2^2 - \pi\left\{\tfrac{1}{2}a_0^2 + \sum_{k=1}^{n} (a_k^2 + b_k^2)\right\}$$
$$+ \pi\left\{\tfrac{1}{2}(\alpha_0 - a_0)^2 + \sum_{k=1}^{n} [(\alpha_k - a_k)^2 + (\beta_k - b_k)^2]\right\}.$$

en donde a_k, b_k denotan a los coeficientes de Fourier de f.

DEMOSTRACION. Se tiene

$$\|f - T_n\|_2^2 = \int_{-\pi}^{\pi} [f(t) - T_n(t)]^2\, dt$$
$$= \int_{-\pi}^{\pi} [f(t)]^2\, dt - 2\int_{-\pi}^{\pi} f(t)T_n(t)\, dt + \int_{-\pi}^{\pi} [T_n(t)]^2\, dt.$$

Ahora se puede ver con facilidad que

$$\int_{-\pi}^{\pi} f(t)T_n(t)\, dt = \tfrac{1}{2}\alpha_0 \int_{-\pi}^{\pi} f(t)\, dt + \sum_{k=1}^{n} \alpha_k \int_{-\pi}^{\pi} f(t) \cos kt\, dt$$
$$+ \sum_{k=1}^{n} \beta_k \int_{-\pi}^{\pi} f(t) \operatorname{sen} kt\, dt$$

$$= \pi\left\{\tfrac{1}{2}\alpha_0 a_0 + \sum_{k=1}^{n} (\alpha_k a_k + \beta_k b_k)\right\}.$$

Más aún, usando las relaciones antes citadas se puede ver que

$$\int_{-\pi}^{\pi} [T_n(t)]^2\, dt = \pi\left\{\tfrac{1}{2}\alpha_0{}^2 + \sum_{k=1}^{n} (\alpha_k{}^2 + \beta_k{}^2)\right\}.$$

Si se insertan estas dos relaciones en la primera fórmula y se suma y resta $\pi\{\frac{1}{2}a_0{}^2 + \sum_{k=1}^{n} (a_k{}^2 + b_k{}^2)\}$, se obtiene la fórmula (38.5). Q.E.D.

El lema 38.3 tiene la siguiente interpretación "geométrica" importante: entre todos los polinomios trigonométicos T_n de grado n el que minimiza la expresión $\|f - T_n\|_2{}^2$ está determinado de manera única y se obtiene escogiendo a los coeficientes α_k, β_k como los coeficientes de Fourier a_k, b_k de f, $k = 0, 1, \ldots, n$. Si se denota a este polinomio trigonométrico minimizante (único) por medio de $S_n(f)$, entonces

(38.6) $$S_n(f)(x) = \tfrac{1}{2}a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \operatorname{sen} kx)$$

es la n-ésima suma parcial de la serie de Fourier para f y la fórmula (38.5) implica

(38.7) $$\|f - S_n(f)\|_2{}^2 = \|f\|_2{}^2 - \pi\left\{\tfrac{1}{2}a_0{}^2 + \sum_{k=1}^{n} (a_k{}^2 + b_k{}^2)\right\}.$$

Haciendo uso del ejercicio 26.F se puede probar que

(38.8) $$\lim_n \|f - S_n(f)\|_2 = 0$$

para cada función continua con periodo 2π. Sin embargo, puesto que ese ejercicio es el resultado de un análisis considerable, es preferible obtener este resultado de una manera más directa. Para hacer esto serán necesarios los dos siguientes resultados.

38.4 DESIGUALDAD DE BESSEL. *Si* $f \in PC(2\pi)$, *entonces*

(38.9) $$\tfrac{1}{2}a_0{}^2 + \sum_{k=1}^{\infty} (a_k{}^2 + b_k{}^2) \le \frac{1}{\pi}\|f\|_2{}^2.$$

DEMOSTRACION. Si $n \in N$ es arbitraria, entonces de (38.7) se deduce que

$$\tfrac{1}{2}a_0{}^2 + \sum_{k=1}^{n} (a_k{}^2 + b_k{}^2) \le \frac{1}{\pi}\|f\|_2{}^2.$$

Por lo tanto, las sumas parciales de la serie del lado izquierdo en (38.9) estan acotadas por arriba. Puesto que los términos son todos positivos, esta serie es convergente y se cumple (38.9). Q.E.D.

El siguiente resultado es un caso especial de lo que por lo general se llama lema de Riemann-Lebesgue.

38.5 LEMA DE RIEMANN-LEBESGUE. *Si* $g \in PC(2\pi)$, *entonces*

$$\lim_n \int_{-\pi}^{\pi} g(t) \operatorname{sen}(n+\tfrac{1}{2})t \, dt = 0.$$

DEMOSTRACION. Dado que $\operatorname{sen}(n+\frac{1}{2})t = \operatorname{sen} nt \cos \frac{1}{2}t + \cos nt \operatorname{sen} \frac{1}{2}t$, se tiene

$$\int_{-\pi}^{\pi} g(t) \operatorname{sen}(n+\tfrac{1}{2})t \, dt = \frac{1}{\pi}\int_{-\pi}^{\pi} [\pi g(t) \cos \tfrac{1}{2}t] \operatorname{sen} nt \, dt + \frac{1}{\pi}\int_{-\pi}^{\pi} [\pi g(t) \operatorname{sen} \tfrac{1}{2}t] \cos nt \, dt.$$

Puesto que $g \in PC(2\pi)$, se deduce que las funciones definidas para $t \in (-\pi, \pi]$, por medio de

$$g_1(t) = \pi g(t) \cos \tfrac{1}{2}t, \qquad g_2(t) = \pi g(t) \operatorname{sen} \tfrac{1}{2}t,$$

tienen extensiones a $\boldsymbol{R}$ que pertenecen a $PC(2\pi)$. Por lo tanto, las integrales del lado derecho de la fórmula anterior dan los coeficientes de Fourier para g_1 y g_2; por lo que por la desigualdad de Bessel estas Integrales convergen a 0 cuando $n \to \infty$. Q.E.D.

38.6 LEMA. *Si* $f \in PC(2\pi)$, *entonces la suma parcial* $S_n(f)$ *de su serie de Fourier está dada por*

$$S_n(f)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t) D_n(t) \, dt \tag{38.10}$$

en donde D_n *es el n-ésimo kernel de Dirichlet definido por*

$$D_n(t) = \tfrac{1}{2} + \sum_{k=1}^{n} \cos kt = \begin{cases} \dfrac{\operatorname{sen}(n+\frac{1}{2})t}{2 \operatorname{sen} \frac{1}{2}t}, & 0 < |t| \leq \pi, \\ n + \frac{1}{2}, & t = 0. \end{cases}$$

DEMOSTRACION. De las fórmulas (38.2) y (38.6) se deduce que

$$\begin{aligned} S_n(f)(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t) \, dt + \frac{1}{\pi}\int_{-\pi}^{\pi} \sum_{k=1}^{n} f(t)\{\cos kx \cos kt + \operatorname{sen} kx \operatorname{sen} kt\} \, dt \\ &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left\{\tfrac{1}{2} + \sum_{k=1}^{n} \cos k(x-t)\right\} dt. \end{aligned}$$

Si se toma $t = x + s$ y se usa el hecho de que el coseno es una función par y que la integral tiene periodo 2π, se tiene

$$S_n(f)(x) = \frac{1}{\pi}\int_{-\pi-x}^{\pi-x} f(x+s)\left\{\tfrac{1}{2} + \sum_{k=1}^{n} \cos ks\right\} ds$$
$$= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+s)\left\{\tfrac{1}{2} + \sum_{k=1}^{n} \cos ks\right\} ds.$$

Se aplica la fórmula (36.6) para obtener (38.10). Q.E.D.

Antes de seguir adelante recuerde (ver el ejercicio 27.Q) que la **derivada por la derecha** de una función $f: \mathbf{R} \to \mathbf{R}$ en un punto $c \in \mathbf{R}$ en donde f tiene por la derecha $f(c+)$, es el límite

$$f'_+(c) = \lim_{\substack{t \to 0 \\ t>0}} \frac{f(c+t) - f(c+)}{t}$$

siempre que este límite exista. De manera análoga, la **derivada por la izquierda** de f en c es el límite.

$$f'_-(c) = \lim_{\substack{t \to 0 \\ t<0}} \frac{f(c+t) - f(c-)}{t}$$

38.7 TEOREMA DE LA CONVERGENCIA PUNTUAL. *Suponga que $f \in PC(2\pi)$ y que f tiene derivadas por la derecha y por la izquierda en c. Entonces, la serie de Fourier para f converge a $\frac{1}{2}\{f(c-)+f(c+)\}$ en el punto c. Expresado en símbolos,*

$$\text{(38.11)} \qquad \tfrac{1}{2}\{f(c-)+f(c+)\} = \tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} (a_n \cos nc + b_n \operatorname{sen} nc).$$

DEMOSTRACION. De (36.6) se deduce que si sen $\frac{1}{2}t \neq 0$, entonces

$$\tfrac{1}{2} + \sum_{k=1}^{n} \cos kt = \frac{\operatorname{sen}(n+\frac{1}{2})t}{2 \operatorname{sen} \frac{1}{2}t}.$$

Multiplicar por $(1/\pi)f(c+)$ e integrar con respecto a t sobre $[0, \pi]$. Dado que $\int_0^\pi \cos kt\, dt = 0$ para $k \in \mathbf{N}$, se obtiene

$$\tfrac{1}{2}f(c+) = \frac{1}{\pi}\int_0^{\pi} f(c+) \frac{\operatorname{sen}(n+\frac{1}{2})t}{2 \operatorname{sen} \frac{1}{2}t}\, dt.$$

En forma análoga, si se multiplica la expresión anterior por $(1/\pi)f(c-)$ y se integra con respecto a t sobre $[-\pi, 0]$, se obtiene

$$\tfrac{1}{2}f(c-) = \frac{1}{\pi}\int_{-\pi}^{0} f(c-) \frac{\operatorname{sen}(n+\frac{1}{2})t}{2 \operatorname{sen} \frac{1}{2}t}\, dt.$$

Si se restan estas expresiones de la fórmula (38.10) se obtiene

$$(*) \qquad S_n(f)(c) - \tfrac{1}{2}\{f(c-)+f(c+)\} = \frac{1}{\pi}\int_{-\pi}^{0} \frac{f(c+t)-f(c-)}{2\,\mathrm{sen}\,\frac{1}{2}t}\,\mathrm{sen}\,(n+\tfrac{1}{2})t\,dt$$
$$+\frac{1}{\pi}\int_{0}^{\pi} \frac{f(c+t)-f(c+)}{2\,\mathrm{sen}\,\frac{1}{2}t}\,\mathrm{sen}(n+\tfrac{1}{2})t\,dt.$$

Ahora, puesto que

$$\lim_{\substack{t\to 0\\ t>0}} \frac{f(c+t)-f(c+)}{2\,\mathrm{sen}\,\frac{1}{2}t} = \lim_{\substack{t\to 0\\ t>0}} \left\{\frac{f(c+t)-f(c+)}{t}\cdot\frac{t}{2\,\mathrm{sen}\,\frac{1}{2}t}\right\}$$
$$= f'_+(c)\cdot 1 = f'_+(c),$$

se infiere que la función

$$\begin{aligned} F_+(t) &= \frac{f(c+t)-f(c+)}{2\,\mathrm{sen}\,\frac{1}{2}t} && \text{para} \quad t\in(0,\pi],\\ &= f'_+(c) && \text{para} \quad t=0,\\ &= 0 && \text{para} \quad t\in(-\pi,0), \end{aligned}$$

es continua por partes en $(-\pi, \pi]$. Por lo tanto, la segunda integral en (*9 converge a 0 cuando $n \to \infty$.

En forma análoga, la primera integral en (*) converge a 0 cuando $n \to \infty$. Por lo tanto, se obtiene la conclusión deseada. Q.E.D.

38.8 EJEMPLOS. *(a)* La función f_1 del ejemplo 38.2(*a*) está en $PC(2\pi)$, con $f(c-) = f(c) = f(c+)$ para $c\in[-\pi, \pi]$, $c \neq -\pi, 0, +\pi$, en donde se tiene $f(-\pi-) = +1$, $f(-\pi+) = -1$, $f(0-) = -1$, $f(0+) = 1$, $f(\pi-) = 1$, en donde $f(\pi+) = -1$. Puesto que las derivadas unilaterales existen en todas partes (y son iguales a 0), se deduce del teorema de convergencia puntual 38.7 que la serie de Fourier para f_1 converge a $f_1(c)$ siempre que $c\in[-\pi, \pi]$, $c\neq -\pi, 0, \pi$ y que en estos tres puntos la serie de Fourier para f_1 converge a 0.

(b) La función f_2 del ejemplo 38.2(*b*) es continua, tiene periodo 2π y tiene derivadas unilaterales en todas partes. Por lo tanto, la serie de Fourier para f_2 converge en todo punto a f_2 y, como ya se vio, la convergencia es uniforme. Se puede ver que la derivada (por ambos lados) de f_2 existe en $[-\pi, \pi]$ excepto en los puntos $0, \pm\pi$ y que f'_2 coincide con la función continua por partes f_1 para $x \notin \{n\pi : n \in \mathbf{Z}\}$.

Recuerde que del teorema del valor medio (véase el ejercicio 27.N) se deduce que si $f' \in PC(2\pi)$, entonces las derivadas de f por la izquierda y por la derecha existen en los puntos de discontinuidad de f'. En seguida se demuestra que para una función f con periodo 2π y tal que $f' \in PC(2\pi)$, la serie de Fourier para f es *uniformemente convergente a* f.

38.9 TEOREMA DE CONVERGENCIA UNIFORME. *Sea f continua, con periodo 2π y supóngase que $f' \in PC(2\pi)$. Entonces, la serie de Fourier para f converge uniformemente a f en* $\mathbf{R}$.

DEMOSTRACION. Dado que f es continua y las derivadas unilaterales de f existen en todo punto, del teorema de convergencia 38.7 se infiere que la serie de Fourier para f converge a f en todo punto. Falta demostrar que la convergencia es uniforme. Dada la desigualdad

$$\left|\sum_{k=1}^{\infty}(a_k \cos kx + b_k \operatorname{sen} kx)\right| \leq \sum_{k=1}^{\infty}(|a_k| + |b_k|),$$

basta con probar la convergencia de la última serie. De hecho, si se aplica la desigualdad de Bessel a f', se sabe que la serie $\sum(|a_k'|^2 + |b_k'|^2)$ es convergente. Pero como se vio en el ejemplo 38.2(e), $a_k = -b_k'/k$ y $b_k = a_k'/k$. Aplicando la desigualdad de Schwarz se tiene

$$\sum_{k=1}^{m}|a_k| = \sum_{k=1}^{m}\frac{1}{k}|b_k'| \leq \left(\sum_{k=1}^{m}\frac{1}{k^2}\right)^{1/2}\left(\sum_{k=1}^{m}|b_k'|^2\right)^{1/2}.$$

Puesto que una desigualdad similar es válida para $\sum|b_k|$, se obtiene la afirmación deseada. Q.E.D.

Se demostrará ahora que las sumas parciales de la serie de Fourier para cualquier función f en $PC(2\pi)$ convergen a f en la norma $\|\cdot\|_2$. Aunque esto no garantiza que se pueda recuperar el valor de f en algún punto en particular asignado con anterioridad, se pueden interpretar como que da f en un cierto sentido "estadístico". Para algunas aplicaciones este tipo de convergencia es tan útil como la convergencia puntual y existe la ventaja de que no se tienen que imponer restricciones de diferenciabilidad.

38.10 TEOREMA DE LA CONVERGENCIA NORMADA. *Si $f \in PC(2\pi)$ y si $(S_n(f))$ es la sucesión de sumas parciales de la serie de Fourier para f, entonces*

$$\lim_{n\to\infty}\|f - S_n(f)\|_2 = 0.$$

DEMOSTRACION. Sea $f \in PC(2\pi)$ y sea $\varepsilon > 0$ dada. Queda como ejercicio demostrar que existe una función continua f_1 con periodo 2π tal que $\|f - f_1\|_2 < \varepsilon/7$. Por el teorema 24.5, existe una función continua lineal por partes f_2 que se puede escoger con periodo 2π, y tal que $\|f_1 - f_2\|_\infty < \varepsilon/7$. Del teorema de convergencia uniforme 38.9 se deduce que si n es suficientemente grande, entonces $\|f_2 - S_n(f_2)\|_\infty < \varepsilon/7$. De la fórmula (38.1) se tiene $\|g\|_2 \leq \sqrt{2\pi}\,\|g\|_\infty \leq 3\,\|g\|_\infty$ para cualquier $g \in PC(2\pi)$; por lo tanto, se deduce que

$$\|f - S_n(f_2)\|_2 < \|f - f_1\|_2 + \|f_1 - f_2\|_2 + \|f_2 - S_n(f_2)\|_2$$

$$\leq \frac{\varepsilon}{7} + \frac{3\varepsilon}{7} + \frac{3\varepsilon}{7} = \varepsilon.$$

Ahora, $S_n(f_2)$ es un polinomio trigonométrico de grado n que aproxima a f en x (con respecto a $\|\cdot\|_2$). Puesto que en el lema 38.3 se probó que la suma parcial $S_n(f)$ es el polinomio trigonométrico de grado n que da la mejor aproximación, se infiere que $\|f - S_n(f)\|_2 < \varepsilon$. Dado que $\varepsilon > 0$ es arbitraria, se concluye que $\lim \|f - S_n(f)\|_2 = 0$. Q.E.D.

El siguiente reforzamiento a la desigualdad de Bessel para $f \in PC(2\pi)$. se obtiene como un corolario de este resultado y del lema 38.3.

38.11 IGUALDAD DE PARSEVAL. *Si* $f \in PC(2\pi)$, *entonces*

$$\frac{1}{\pi}\|f\|_2^2 = \tfrac{1}{2}a_0^2 + \sum_{k=1}^{\infty}(a_k^2 + b_k^2), \tag{38.12}$$

en donde a_k, b_k *son los coeficientes de Fourier de* f.

Esta sección se terminará con una demostración del teorema de Fejér† acerca de la sumabilidad de Cesâro de la serie de Fourier de una función continua. Si $S_n(f)$, $n = 0, 1, 2, \ldots$, designa las sumas parciales de la serie de Fourier correspondiente a f, desígnese con $\Gamma_n(f)$ la medida de Cesâro:

$$\Gamma_n(f) = \frac{1}{n}[S_0(f) + S_1(f) + \cdots + S_{n-1}(f)].$$

Sea D_n, $n = 0, 1, 2, \ldots$, como en el lema 38.6. Haciendo uso de la fórmula elemental

$$2\,\text{sen}\,(k - \tfrac{1}{2})t\,\text{sen}\,\tfrac{1}{2}t = \cos\,(k-1)t - \cos kt,\ k = 0, 1, 2, \ldots,$$

se puede probar que

$$\frac{1}{n}[D_0(t) + D_1(t) + \cdots + D_{n-1}(t)] = \begin{cases} \dfrac{1}{2n}\left(\dfrac{\text{sen}\,\frac{1}{2}nt}{\text{sen}\,\frac{1}{2}t}\right)^2, & 0 < |t| \le \pi, \\ \tfrac{1}{2}n, & t = 0, \end{cases} \tag{38.13}$$

y K_n va a ser esta función a la que se va a llamar el n-ésimo kernel de Fejér. Es claro que $K_n(t) \ge 0$ y puesto que

$$\frac{1}{\pi}\int_{-\pi}^{\pi} D_k(t)\,dt = 1$$

para $k = 0, 1, 2, \ldots$, se deduce que

$$\frac{1}{\pi}\int_{-\pi}^{\pi} K_n(t)\,dt = 1. \tag{38.14}$$

† LEOPOLD FEJER (1880-1959) estudió y dio clases en Budapest. Hizo muchas aportaciones interesantes en varias áreas de análisis real y complejo.

Además, si $0<\delta<\pi$, del hecho de que sen $\theta \geq 2\theta/\pi$ para $0 \leq \theta \leq \pi/2$ se deduce que

(38.15) $$0 \leq K_n(t) \leq \frac{1}{2n}\left(\frac{\pi}{2\delta}\right)^2 \quad \text{para } \delta \leq |t| \leq \pi.$$

Por último, observe que por el lema 38.6 se puede expresar la medida de Cesâro por medio de la fórmula

(38.16) $$\Gamma_n(f)(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)K_n(t)\,dt.$$

Se puede pasar ahora a la demostración del teorema de Fejer.

38.12 TEOREMA DE FEJER. *Si f es continua y tiene periodo 2π, entonces la media de Cesâro de la serie de Fourier para f converge uniformemente a f en R.*

DEMOSTRACION. De (38.14) se infiere que

$$f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)K_n(t)\,dt.$$

Restando esto de (38.16) se obtiene

$$\Gamma_n(f)(x) - f(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} \{f(x+t)-f(x)\}K_n(t)\,dt.$$

Dado que $K_n(t) \geq 0$ para toda t, se tiene

$$|\Gamma_n(f)(x) - f(x)| \leq \frac{1}{\pi}\int_{-\pi}^{\pi} |f(x+t)-f(x)|\,K_n(t)\,dt.$$

Sea $\varepsilon > 0$ dada, ya que f es uniformemente continua en $\boldsymbol{R}$, existe un número δ con $0<\delta<\pi$ tal que si $|t| \leq \delta$, entonces

$$|f(x+t)-f(x)| < \varepsilon \qquad \text{para toda } x \in [-\pi, \pi].$$

Por lo tanto, se tiene

$$\frac{1}{\pi}\int_{-\delta}^{\delta} |f(x+t)-f(x)|\,K_n(t)\,dt \leq \frac{\varepsilon}{\pi}\int_{-\delta}^{\delta} K_n(t)\,dt \leq \frac{\varepsilon}{\pi}\int_{-\pi}^{\pi} K_n(t) = \frac{\varepsilon}{\pi}.$$

Por otro lado, por (38.15) se tiene

$$\frac{1}{\pi}\int_{\delta}^{\pi} |f(x+t)-f(x)|\,K_n(t)\,dt \leq \frac{\pi-\delta}{\pi}(2\,\|f\|_\infty)\left(\frac{1}{8n}\frac{\pi^2}{\delta^2}\right) \leq \frac{1}{n}\left(\frac{\pi^2\|f\|_\infty}{4\delta^2}\right),$$

que se puede hacer menor que ε tomando a n suficientemente grande. Puesto que una estima análoga es válida para la integral en $[-\pi, -\delta]$, se infiere que

$$\|\Gamma_n(f)-f\|_\infty<\left(2+\frac{1}{\pi}\right)\varepsilon$$

para n suficientemente grande Q.E.D.

Puesto que se puede ver con facilidad que la función $\Gamma_n(f)$ es un polinomio trigonométrico (de grado $n-1$), se tiene otra demostración para el siguiente teorema de Weierstrass.

38.13 TEOREMA DE APROXIMACION DE WEIERSTRASS. *Si f es continua y tiene periodo 2π, entonces se puede aproximar uniformemente por polinomios trigonométricos.*

Ejercicios

38.A. Sea g una función de valor real definida en una celda J en $\boldsymbol{R}$ con puntos extremos $a<b$. Se dice que g es **continua por partes** en J si (i) g tiene un límite por la derecha y por la izquierda.

(a) Demostrar que si g es continua por partes en $(-\pi, \pi]$, entonces existe una función única G en $PC(2\pi)$ tal que $G(x)=g(x)$ para toda $x\in(-\pi,\pi]$.

(b) La función g tiene una derivada bilateral por la izquierda (respectivamente, por la derecha) en $c\in(-\pi,\pi)$ si y sólo si G la tiene.

(c) La función g tiene una derivada por la derecha en $-\pi$ (respectivamente, una derivada por la izquierda en π) si y sólo si G la tiene.

(d) Las derivadas unilaterales $g_+'(-\pi)$, $g_-'(\pi)$ existen y son iguales si y sólo si G tiene una derivada en $\pm\pi$.

38.B. Si $f\in PC(2\pi)$ y la derivada $f'(x)$ existe para toda $x\in\boldsymbol{R}$, entonces f' tiene periodo 2π.

(b) Si $f\in PC(2\pi)$ y $c\in\boldsymbol{R}$, defina $F:\boldsymbol{R}\to\boldsymbol{R}$ por medio de $F(x)=\int_c^x f(t)\,dt$, de tal manera que F sea continua. Demostrar que F tiene periodo 2π si y sólo si la **media** de f es cero, es decir

$$\frac{1}{2}a_0=\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\,dt=0.$$

38.C. *(a)* Sea $f\in PC(2\pi)$ impar. Entonces, $f(\pm\pi)=0$. Si f es continua en 0, entonces $f(0)=0$.

(b) Sea $g\in PC(2\pi)$ par; entonces $g(0+)=g(0-)$. Si la derivada $g'(x)$ existe para toda $x\in\boldsymbol{R}$, entonces (véase el ejercicio 27.P) g' es impar, tiene periodo 2π y $g'(0)=g'(\pm\pi)=0$.

38.D. Sean F y f funciones de $PC(2\pi)$ que tienen coeficientes de Fourier A_n, B_n y a_n, b_n, respectivamente. Si $\alpha,\beta\in\boldsymbol{R}$ y si $h=\alpha F+\beta f$, demostrar que h pertenece a $PC(2\pi)$ y tiene coeficientes de Fourier $\alpha A_n+\beta a_n$, $\alpha B_n+\beta b_n$. (de donde los coeficientes de Fourier de una función dependen linealmente en la función).

38.E. *(a)* Sea f_1 la función del ejemplo 38.2(*a*). Calcular la serie de Fourier para f_1 y demostrar que esta serie de Fourier no converge uniformemente en $[-\pi,\pi]$.

(b) Sea f_2 la función del ejemplo 38.2(*b*). Calcular la serie de Fourier para f_2 y demostrar que la derivada término por término de la serie de Fourier para f_2 coincide con la serie de Fourier para f_1.

(c) Usando el hecho de que la serie de Fourier para f_2 converge a f_2, deducir que

$$\frac{\pi^2}{8}=\frac{1}{1^2}+\frac{1}{3^2}+\frac{1}{5^2}+\cdots.$$

(d) Sea $f_3(x)=\frac{1}{2}\pi-f_2(x)$ de tal manera que $f_3(x)=\frac{1}{2}\pi-|x|$ para $x\in(-\pi,\pi]$. Usar el ejercicio 38.D para demostrar que la serie de Fourier para f_3 está dada por

$$f_3(x)\sim\frac{4}{\pi}\left[\frac{\cos x}{1^2}+\frac{\cos 3x}{3^2}+\frac{\cos 5x}{5^2}+\cdots\right].$$

38.F. *(a)* Sea $g_1\in PC(2\pi)$ tal que $g_1(x)=x$ para $x\in(-\pi,\pi]$ y $g_1(\pi)=0$. Demostrar que g_1 es una función impar y que la serie de Fourier está dada por

$$2\left[\frac{\operatorname{sen} x}{1}-\frac{\operatorname{sen} 2x}{2}+\frac{\operatorname{sen} 3x}{3}-\cdots\right].$$

Obsérvese que esta serie de Fourier converge a 0 en $x=\pm\pi$. Usar el teorema de convergencia puntual 38.7 para demostrar que esta serie de Fourier converge a $g_1(x)$ para todo punto $x\in[-\pi,\pi]$.

(b) Sea $g_2\in PC(2\pi)$ tal que $g_2(x)=x^2$ para $x\in(-\pi,\pi]$.. Demostrar que g_2 es una función par y que su serie de Fourier está dada por

$$\frac{\pi^2}{3}-4\left[\frac{\cos x}{1^2}-\frac{\cos 2x}{2^2}+\frac{\cos 3x}{3^2}-\cdots\right].$$

Demostrar que esta serie de Fourier converge uniformemente a g_2 en $[-\pi,\pi]$ y que su derivada término por término es el doble de la serie de Fourier para g_1.

(c) Demostrar que

$$\frac{\pi^2}{12}=\frac{1}{1^2}-\frac{1}{2^2}+\frac{1}{3^2}-\cdots.$$

(d) Sea $h(x)=\frac{1}{3}\pi^2-g_2(x)$ tal que $h(x)=\frac{1}{3}\pi^2-x^2$ para $x\in(-\pi,\pi]$. Entonces, la serie de Fourier para h está dada por

$$4\left[\frac{\cos x}{1^2}-\frac{\cos 2x}{2^2}+\frac{\cos 3x}{3^2}-\cdots\right].$$

38.G. *(a)* Sea $k(x)=x^3$ para toda $x\in\boldsymbol{R}$. Demostrar que k es continua e impar en $\boldsymbol{R}$. Sin embargo, la función k_1 en $PC(2\pi)$ que coincide con K en $(-\pi,\pi]$ no es continua.

(b) Sea $h(x)=x^3-\pi^2x$ se tiene que h es continua e impar en $\boldsymbol{R}$. Sea h_1 la función en $PC(2\pi)$ que coincide con h en $(-\pi,\pi]$. Demostrar que h_1 es continua en $\boldsymbol{R}$ y que $h_1'(x)=3x^2-\pi^2$ para $x\in(-\pi,\pi]$.

(c) Usar el ejercicio 87.P, el ejemplo 38.2(*e*) y el ejercicio 38.C(*d*) para demostrar que la serie de Fourier para h_1 está dada por

$$-12\left[\frac{\operatorname{sen} x}{1^3}-\frac{\operatorname{sen} 2x}{2^3}+\frac{\operatorname{sen} 3x}{3^2}-\cdots\right].$$

38.H. Sea $f:[0, \pi] \to \mathbf{R}$ continua por partes y sea $f_e \in PC(2\pi)$ la función definida por

$$f_e(x) = f(x) \quad \text{para } x \in [0, \pi],$$
$$= f(-x) \quad \text{para } x \in [-\pi, 0).$$

(a) Demostrar que f_e es una función par, se le llama **extensión par** de f con periodo 2π.

(b) A la serie de Fourier de f_e se la llama serie coseno (de Fourier) de f. Demostrar que está dada por

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty} a_n \cos nx.$$

en donde

$$a_n = \frac{2}{\pi}\int_0^{\pi} f(t)\cos nt\, dt, \qquad n = 0, 1, 2, \ldots,$$

(c) Demostrar que si $c \in (0, \pi)$ y f tiene derivadas por la izquierda y por la derecha en c, entonces la serie coseno para f converge a $\tfrac{1}{2}[f(c-) + f(c+)]$. Asimismo, si f tiene una derivada por la derecha en 0, entonces la serie coseno para f converge a $f(0+)$. Si f tien una derivada por la izquierda en π, entonces la serie coseno para f converge a $f(\pi-)$.

38.I. Para cada una de las siguientes funciones definidas en $[0, \pi]$, calcular la serie coseno y determinar el límite de esta serie en cada punto.

(a) $f(x) = x$;

(b) $f(x) = \operatorname{sen} x$;

(c) $f(x) = 1$ para $0 \le x \le \tfrac{1}{2}\pi$,
$= 0$ para $\tfrac{1}{2}\pi < x \le \pi$.

(d) $f(x) = \tfrac{1}{2}\pi - x$ para $0 \le x \le \tfrac{1}{2}\pi$,
$= 0$ para $\tfrac{1}{2}\pi < x \le \pi$.

(e) $f(x) = x(\pi - x)$.

38.J. Sea $f:[0, \pi] \to \mathbf{R}$ continua por partes y defínase $f_o \in PC(2\pi)$ por medio de

$$f_o = f(x) \quad \text{para } x \in (0, \pi],$$
$$= 0 \quad \text{para } x = 0,$$
$$= -f(-x) \quad \text{para } x \in (-\pi, 0).$$

(a) Demostrar que f_o es una función impar, se la llama **extensión impar** de f con periodo 2π.

(b) La serie de Fourier de f_o se llama **serie seno** (de Fourier) de f. Demostrar que está dada por

$$\sum_{n=1}^{\infty} b_n \operatorname{sen} nx,$$

en donde

$$b_2 = \frac{2}{\pi}\int_0^{\pi} f(t)\operatorname{sen} nt\, dt, \qquad n = 1, 2, \ldots.$$

(c) Demostrar que si $c \in (0, \pi)$ y si f tiene derivadas por la izquierda y por la derecha en c, entonces la serie seno para f converge a $\frac{1}{2}[f(c-)+f(c+)]$. En cualquiera de los casos la serie seno para f converge a 0 en $x=0, \pi$.

38.K. Para cada una de las siguientes funciones definidas en $[0, \pi]$, calcular la serie seno y determinar el límite de esta serie en cada punto.

(a) $f(x)=1$;

(b) $f(x)=\cos x$;

(c) $f(x)=1$ para $0 \le x \le \frac{1}{2}\pi$,
$=0$ para $\frac{1}{2}\pi < x \le \pi$;

(d) $f(x)=\pi - x$;

(e) $f(x)=x(\pi-x)$.

38.L. Sea $f_n \in PC(2\pi)$ la función tal que $f_n(x)=n^{1/4}$ para $0 \le x \le 1/n$ $=0$ para otra $x \in (-\pi, \pi]$. Demostrar que $\|f_n\|_2 = 1/n^{1/4}$ de tal manera que la sucesión (f_n) converge a la función cero en la norma $\|\cdot\|_2$ pero dado que es no acotada, la convergencia no es uniforme.

38.M. Si $f \in PC(2\pi)$ y si $\varepsilon>0$, demostrar que existe una función continua f_1 con periodo 2π tal que $\|f-f_1\|_2<\varepsilon$.

38.N. Mediante la igualdad de Parseval 38.11 compruebe las siguientes fórmulas

(a) $\dfrac{\pi^2}{6}=\sum_{n=1}^{\infty}\dfrac{1}{n^2}$,

(b) $\dfrac{\pi^2}{8}=\sum_{n=1}^{\infty}\dfrac{1}{(2n-1)^2}$,

(c) $\dfrac{\pi^4}{90}=\sum_{n=1}^{\infty}\dfrac{1}{n^4}$,

(d) $\dfrac{\pi^6}{945}=\sum_{n=1}^{\infty}\dfrac{1}{n^6}$.

38.O. Si f y F pertenecen a $PC(2\pi)$ y tienen coeficientes de Fourier a_n, b_n y A_n, B_n, respectivamente, demostrar que

$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)F(t)\,dt = \tfrac{1}{2}a_0A_0 + \sum_{n=1}^{\infty}(a_nA_n+b_nB_n).$$

(Sugerencia: aplicar la igualdad de Parseval a $f+F$.)

38.P. Usar la prueba de Dirichlet 36.2 y el ejemplo 36.8 para demostrar que la serie trigonométrica

$$\sum_{n=1}^{\infty}\frac{\operatorname{sen} nx}{n^{1/2}}$$

converge para toda x. Demostrar que sin embargo esta serie no puede ser la serie de Fourier de ninguna función en $PC(2\pi)$.

38.Q. Sea $L>0$ y sea $PC(2L)$ el espacio vectorial de todas las funciones $f:\mathbf{R}\to\mathbf{R}$ que tienen periodo 2L y que son continuas por partes.

(a) Si se define $f\cdot g=\int_{-L}^{L} f(t)g(t)\,dt$ para f, $g \in PC(2L)$, demostrar que la aplicación $(f, g)\mapsto f\cdot g$ es un producto interno (en el sentido de la definición 8.3) en $PC(2L)$. Más aún, la norma inducida por este producto interno (véase 8.7) es

$$\|f\|_2=\left[\int_{-L}^{L}|f(t)|^2\,dt\right]^{1/2}.$$

(b) Se van a tomar C_0, C_n, S_n, $n\in\mathbf{N}$, como las funciones en $PC(2L)$ dadas por

$$C_0(x)=\frac{1}{\sqrt{2L}},\qquad C_n(x)=\frac{1}{\sqrt{L}}\cos\frac{n\pi x}{L},\qquad S_n(x)=\frac{1}{\sqrt{L}}\operatorname{sen}\frac{n\pi x}{L}.$$

Demostrar que este conjunto de funciones es **ortonormal** en el sentido de que

$$C_n \cdot S_m = 0, \qquad C_n \cdot C_m = \delta_{nm}, \qquad S_n \cdot S_m = \delta_{nm}$$

en donde $\delta_{nm} = 1 \quad n = m$ y $\delta_{nm} = 0$ si $n \neq m$. (*Sugerencia:* si $L = \pi$, éstas son las relaciones dadas antes de 38.3.)

(c) Si $f \in PC(2L)$, se define la **serie de Fourier** de f en $[-L, L]$ como la serie

$$\frac{1}{2} a_0 + \sum_{n=1}^{\infty} \left(a_n \cos \frac{n\pi x}{L} + b_n \operatorname{sen} \frac{n\pi x}{L} \right),$$

en donde se tiene

$$a_0 = \frac{1}{L} \int_{-L}^{L} f(t)\, dt, \qquad a_n = \frac{1}{L} \int_{-L}^{L} f(t) \cos \frac{n\pi t}{L}\, dt,$$

$$b_n = \frac{1}{L} \int_{-L}^{L} f(t) \operatorname{sen} \frac{nt}{L}\, dt,$$

para $n = 1, 2, \ldots$.

(d) Reformular los teoremas de convergencia 38.7, 38.9 y 38.10 para series de Fourier de funciones en $PC(2L)$. (*Sugerencia:* hacer un cambio de variable.)

(e) Si $f \in PC(2L)$, entonces la igualdad de Parseval se convierte en

$$\frac{1}{L} \|f\|_2{}^2 = \frac{1}{2} a_0{}^2 + \sum_{n=1}^{\infty} (a_n{}^2 + b_n{}^2),$$

en donde la norma de f es como en el inciso *(a)* y los coeficientes de Fourier son como en el inciso *(c)*.

38.R. Para cada una de las siguientes funciones en el intervalo especificado calcular la serie de Fourier en este intervalo y determinar el límite de esta serie en cada punto.

(a)	$f(x) = x$	en	$(-2, 2]$;
(b)	$f(x) = 0$	para	$-4 < x < 0$,
	$= x$	para	$0 \leq x \leq 4$;
(c)	$f(x) = 0$	para	$-3 < x < 0$,
	$= 1$	para	$0 \leq x \leq 1$,
	$= 0$	para	$1 < x \leq 3$.

38.S. Sea f continua y con periodo 2π. Demostrar que si la serie de Fourier f converge en $c \in [-\pi, \pi]$ a algún número, entonces converge a $f(c)$.

38.T. Sea f en $PC(2\pi)$ y suponga que $c \in [-\pi, \pi]$. Si $\Gamma_n(f)$ designa la n-ésima media de Fejér definida en (38.16), demostrar que

$$\lim \Gamma_n(f)(c) = \frac{1}{2} [f(c-) + f(c+)].$$

38 U Suponga que f y f' son continuas con periodo 2π y que $f'' \in PC(2\pi)$. *(a)* Demostrar que los coeficientes de Fourier a_n, b_n de f son tales que la serie

$$\sum_{n=1}^{\infty} n^2(|a_n| + |b_n|)$$

es convergente. Por lo tanto, existe una constante $M>0$ tal que $|a_n| \leq M/n^2$ y $|b_n| \leq M/n^2$ para toda $n \in \mathbf{N}$.

(b) Demostrar que la serie de Fourier para f' es la derivada término por término de la serie de Fourier para f.

38.V. *(a)* Si $k \in PC(2\pi)$ y si x_0, $x \in [-\pi, \pi]$, usar la desigualdad de Schwarz para demostrar que

$$\left| \int_{x_0}^{x} k(t)\, dt \right| \leq \|k\|_2 \, |x - x_0|^{1/2} \leq \|k\|_2 \sqrt{2\pi}.$$

(b) Usar el inciso *(a)* y el teorema de convergencia normada 38.10 para demostrar que si $f \in PC(2\pi)$ y $x_0 \in [-\pi, \pi]$, entonces la serie de Fourier para f se puede integrar término por término:

$$\int_{x_0}^{x} f(t)\, dt = \frac{1}{2} a_0(x - x_0) + \sum_{n=1}^{\infty} \int_{x_0}^{x} (a_n \cos nt + b_n \operatorname{sen} nt)\, dt,$$

y la serie que resulta es uniformemente convergente para $x \in [-\pi, \pi]$.

38.W. *(a)* Suponga que $\alpha > 0$ no es un entero. Demostrar que

$$\cos \alpha x = \frac{2\alpha \operatorname{sen} \alpha\pi}{\pi} \left[\frac{1}{2\alpha^2} - \frac{\cos x}{\alpha^2 - 1^2} + \frac{\cos 2x}{\alpha^2 - 2^2} - \frac{\cos 3x}{\alpha^2 - 3^2} + \cdots \right]$$

para toda $x \in [-\pi, \pi]$.

(b) Usar el inciso *(a)* para demostrar que si $x \notin \mathbf{Z}$, entonces

$$\cot \pi x = \frac{1}{\pi x} + \frac{2x}{\pi} \sum_{n=1}^{\infty} \frac{1}{x^2 - n^2},$$

$$\csc \pi x = \frac{1}{\pi x} + \frac{2x}{\pi} \sum_{n=1}^{\infty} \frac{(-1)^n}{x^2 - n^2}.$$

(c) Diferenciar la primera serie en *(b)* término por término (justificar esto) para demostrar que si $x \notin \mathbf{Z}$, entonces

$$\frac{\pi^2}{(\operatorname{sen} \pi x)^2} = \lim_m \sum_{n=-m}^{m} \frac{1}{(x-n)^2}.$$

(d) Integrar la primera serie en *(b)* término por término (justificar esto) para demostrar que si $x \notin \mathbf{Z}$, entonces

$$\frac{\operatorname{sen} \pi x}{\pi x} = \lim_m \left[\left(1 - \frac{x^2}{1^2}\right)\left(1 - \frac{x^2}{2^2}\right) \cdots \left(1 - \frac{x^2}{m^2}\right) \right].$$

VII DIFERENCIACION EN $\boldsymbol{R}^p$

En este capítulo se estudia la teoría de funciones diferenciables en $\boldsymbol{R}^p$ en donde $p > 1$. La teoría es parecida a la que se da en las secciones 27 y 28, pero surgen algunas complicaciones y aspectos nuevos. Varias de estas complicaciones se deben sólo a la inevitable complejidad de la notacion, pero otras surgen porque es posible llegar a un punto $c \in \boldsymbol{R}^p$ desde "muchas direcciones", de tal manera que pueden ocurrir fenómenos nuevos.

En la sección 27 se definió la derivada de una función $f: \boldsymbol{R} \to \boldsymbol{R}$ en un punto $c \in \boldsymbol{R}$ de la manera tradicional, es decir, como el número $L \in \boldsymbol{R}$ tal que

$$L = \lim_{x \to c} \frac{f(x) - f(c)}{x - c},$$

cuando este límite existe. De manera equivalente se pudo haber definido esta derivada como el número L tal que

$$\lim_{x \to c} \frac{|f(x) - f(c) - L(x - c)|}{|x - c|} = 0.$$

Se puede considerar que dicha relación de límite hace preciso el sentido en el que se aproximan los valores $f(x)$, para x suficientemente cerca de c, por medio de los valores de una aplicación afín†

$$x \mapsto f(c) + L(x - c),$$

cuya gráfica da la recta tangente a la gráfica de f en el punto $(c, f(c))$.

En este acceso a la derivada el que se usará para funciones de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$. De manera que la derivada de una función f definida en una vecindad de un punto $c \in \boldsymbol{R}^p$ con valores en $\boldsymbol{R}^q$ será una aplicación lineal $L: \boldsymbol{R}^p \to \boldsymbol{R}^q$ tal que

† En cursos elementales, dicha transformación se denomina "lineal". Sin embargo, para que haya consistencia con el uso más restringido del término "lineal" introducido en la sección 21 se usará el término "afín" al hacer referencia a la función que se obtiene al sumar una constante a una función lineal.

$$\lim_{x \to c} \frac{\|f(x) - f(c) - L(x-c)\|}{\|x-c\|} = 0.$$

De donde se está aproximando $f(x)$, para x suficientemente cerca de c, por medio de la aplicación afín

$$x \mapsto f(c) + L(x-c)$$

de $\boldsymbol{R}^p$ hacia $\boldsymbol{R}^q$. [El lector deberá observar que si $p = 1$, entonces la notación $L(x-c)$ representa el producto de los números reales L y $x - c$; sin embargo, si $p > 1$, entonces $L(x-c)$ denota el valor de la aplicación lineal L en el vector $x - c$.]

En la sección 39 se estudia la definición y se relaciona la derivada con las diversas derivadas "parciales". En la sección 40 se obtienen la regla de la cadena y el teorema del valor medio que son de suma importancia. En la sección 41 se da un análisis detallado de las propiedades de aplicación de funciones diferenciales, llegando a los importantes teoremas de inversión y de funciones implícitas, y culminando con los teoremas de parametrización y rango. La última sección trata de las propiedades extremas de funciones de valor real en $\boldsymbol{R}^p$.

Sección 39 La derivada en $\boldsymbol{R}^p$

En la sección 27 se trató la derivada de una función con dominio y rango en $\boldsymbol{R}$. En esta sección se habrá de considerar una función definida en un subconjunto de $\boldsymbol{R}^p$ y con valores en $\boldsymbol{R}^q$ desde un punto de vista análogo.

Si el lector revisa la definición 27.1 podrá ver que también es aplicable a una función definida en un intervalo J en $\boldsymbol{R}$ y con valores en el espacio cartesiano $\boldsymbol{R}^q$. Desde luego, en este caso L es un vector en $\boldsymbol{R}^q$. El único cambio que se requiere para esta extensión es el de reemplazar el valor absoluto de la ecuación (27.1) por la norma en el espacio $\boldsymbol{R}^q$. Excepto por ello, la definición 27.1 es aplicable al pie de la letra a esta situación más general. El hecho de que valga la pena estudiar esta situación es claro al saber que una función f de J a $\boldsymbol{R}^q$ se puede considerar como una *curva* en el espacio $\boldsymbol{R}^q$ y que la derivada (cuando existe) de esta función en el punto $x = c$ da un *vector tangente* a la curva en el punto $f(c)$. De manera alternativa, si x designa tiempo, entonces la función f es la trayectoria de un punto en $\boldsymbol{R}^q$ y la derivada $f'(c)$ denota al *vector velocidad* del punto en el tiempo $x = c$.

Una investigación más profunda de estas ideas nos haría estudiar geometría diferencial y dinámica más de lo que es deseable en este momento. La meta que se tiene es más concreta: se desea organizar la maquinaria analítica que haría posible una investigación satisfactoria y eliminar la restricción de que el dominio esté en un espacio unidimensional haciendo posible que el dominio pertenezca al espacio cartesiano $\boldsymbol{R}^p$. En seguida se hará esto.

Un análisis de la definición 27.1 prueba que el único lugar en que es necesario que el dominio conste de un subconjunto de $\boldsymbol{R}$ es en la ecuación (27.1),

en donde aparece un cociente. Dado que no tiene ningún significado el cociente de un vector en $\boldsymbol{R}^q$ por un vector en $\boldsymbol{R}^p$, no se puede interpretar la ecuación (27.1) tal como aparece. Por tanto, hay la necesidad de encontrar reformulaciones para esta ecuación. Una posibilidad, que es de gran interés consiste en tomar "rebanadas" que pasen por el punto c en el dominio. Para hacerlo más sencillo, se habrá de suponer que c es un punto interior del dominio D de la función; entonces, para cualquier u en $\boldsymbol{R}^p$, el punto $c+tu$ pertenece a D para números reales t suficientemente pequeños.

39.1 DEFINICION. Sea f la función definida en un subconjunto A de $\boldsymbol{R}^p$ y con valores en $\boldsymbol{R}^q$, sea c un punto interior de A y sea u cualquier punto en $\boldsymbol{R}^p$. Se dice que un vector $L_u \in \boldsymbol{R}^q$ es la **derivada parcial de** f **en** c **con respecto a** u si para cada número $\varepsilon > 0$ hay una $\delta(\varepsilon) > 0$ tal que para toda $t \in \boldsymbol{R}$ que satisfaga $0 < |t| < \delta(\varepsilon)$, se tiene

$$\left\| \frac{1}{t}\{f(c+tu) - f(c)\} - L_u \right\| < \varepsilon. \tag{39.1}$$

Se puede ver fácilmente que la derivada parcial L_u definida en (39.1) está determinada de manera única cuando existe. En forma alternativa, se puede definir L_u como el límite

$$\lim_{t\to 0} \frac{1}{t}\{f(c+tu) - f(c)\},$$

o como la derivada en $t = 0$ de la función F definida por $F(t) = f(c+tu)$ para $|t|$ suficientemente pequeña y con valores en $\boldsymbol{R}^q$.

Se escribirá $D_u f(c)$ o $f_u(c)$ para denotar la derivada parcial L_u de f en c con respecto a u. Es preferible usar la primera notación cuando, como ocurre a menudo, el símbolo que denota a la función tiene un subíndice. Se denota a la función $c \mapsto D_u f(c) = f_u(c)$ por medio de $D_u f$ o f_u; eitá definida para aquellos puntos interiores c de A para los cuales el límite requerido existe y tiene valores en $\boldsymbol{R}^q$.

Es claro que si f es de valor real (de tal manera que $q = 1$) y si u es el vector $e_1 = (1, 0, \ldots, 0)$ en $\boldsymbol{R}^p$, entonces la derivada parcial de f con respecto a e_1 coincide con lo que normalmente se llama la **derivada parcial de** f **con respecto a su primera variable,** que con frecuencia se denota por medio de

$$D_1 f, \qquad f_{x_1}, \qquad \text{ó} \qquad \frac{\partial f}{\partial x_1}.$$

De la misma manera, tomando $e_2 = (0, 1, \ldots, 0), \ldots, e_p = (0, 0, \ldots, 1)$, se obtienen las **derivadas parciales de** f **con respecto a las otras variables.**

$$D_2 f = f_{x_2} = \frac{\partial f}{\partial x_2}, \ldots, D_p f = f_{x_p} = \frac{\partial f}{\partial x_p}.$$

En caso de que el símbolo que denota a la función tenga un subíndice, a menudo se insertará una coma para indicar una derivada parcial, de manera que $D_j f_2 = f_{2,j}$.

Se debe observar que puede existir la derivada parcial de una función en un punto con respecto a un vector; sin embargo la derivada parcial con respecto a otro vector no necesariamente existe (véase el ejercicio 39 A). También es claro que con hipótesis apropiadas hay relaciones algebraicas entre derivadas parciales de sumas y productos de funciones, etc. No se obtendrán aquí estas relaciones ya que son casos especiales de lo que se hará en seguida, o bien se pueden demostrar de manera análoga.

Cabe mencionar algo acerca de la terminología. Si u es un vector unitario en $\boldsymbol{R}^p$, entonces la derivada parcial $D_u f(c) = f_u(c)$ a menudo se llama la **derivada direccional de** f **en** c **en la dirección de** u.

La derivada

El principal inconveniente de la derivada parcial de una función f en un punto c con respecto a un vector u es que sólo da una visión del comportamiento de f cerca de c en el conjunto inidimensional $\{c+tu : t \in \boldsymbol{R}\}$. Para poder obtener una información más completa acerca de f en una vecindad de $c \in \boldsymbol{R}^p$, se introducirá el concepto de la derivada de f en c, que es una *aplicación lineal* de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$.

39.2 DEFINICION. Sea f con dominio A en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$, y sea c un punto interior de A. Se dice que f es **diferenciable** en c si existe una función lineal $L : \boldsymbol{R}^p \to \boldsymbol{R}^q$ tal que paia toda $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que si $x \in \boldsymbol{R}^p$ es cualquier vector que satisface $\|x-c\| \le \delta(\varepsilon)$, entonces $x \in A$

$$\|f(x) - f(c) - L(x-c)\| \le \varepsilon \, \|x-c\|. \tag{39.2}$$

En forma alternativa, (39.2) se puede reformular pidiendo que para cualquier $\varepsilon > 0$ exista $\delta(\varepsilon) > 0$ tal que si $u \in \boldsymbol{R}^p$ y $\|u\| \le \delta(\varepsilon)$, entonces

$$\|f(c+u) - f(c) - L(u)\| \le \varepsilon \, \|u\|, \tag{39.3}$$

que a su vez se puede expresar de una manera más compacta escribiendo

$$\lim_{\|u\| \to 0} \frac{\|f(c+u) - f(c) - L(u)\|}{\|u\|} = 0. \tag{39.4}$$ †

En seguida se verá que dicha función lineal L está determinada de manera única cuando existe. Se le llama la **derivada de** f **en** c y a menudo se denotará por medio de $Df(c)$ en vez de L. Con frecuencia se escribirá $Df(c)(u)$ para denotar $L(u)$ y $Df(c)(x-c)$ para denotar $L(x-c)$.

Desde un punto de vista analítico, la existencia de la derivada de f en c refleja la posibilidad de aproximar la aplicación $x \mapsto f(x)$ por la aplicación $x \mapsto f(c) + L(x-c)$. La desigualdad (39.2) da una medida de la cercanía de

† Se advierte al lector que L algunas veces se llama la **derivada de Fréchet** o la **diferencial** de f en c y algunas veces se denota por $df(c)$ o $f'(c)$, etc.

esta aproximación cuando x está cerca de c. Debido a la linealidad de L se tiene

$$f(c)+L(x-c)=(f(c)-L(c))+L(x).$$

De modo que se está aproximando $x \mapsto f(x)$ una función de la forma $x \mapsto y_0+L(x)$, en donde y_0 es fijo. A dichas funciones se les llama **aplicaciones afines** de $\mathbf{R}^p$ a $\mathbf{R}^q$; son simplemente translaciones de aplicaciónes lineales y por ello son muy sencillas.

Desde un punto de vista geométrico, la existencia de la derivada de f en c refleja la existencia de un **plano tangente** a la superficie $\{(x, f(x)) : x \in A\}$ en $\mathbf{R}^p \times \mathbf{R}^q$ en el punto $(c, f(c))$; específicamente el plano dado por la gráfica

$$\{(x, f(c)+L(x-c)) : x \in \mathbf{R}^p\}. \tag{39.5}$$

Se probará en seguida la unicidad de la derivada.

39.3 LEMA. *La función f tiene cuando mucho una derivada en un punto.*

DEMOSTRACION. Suponga que L_1, L_2 son funciones lineales de $\mathbf{R}^p$ a $\mathbf{R}^q$ y que satisfacen (39.3) para $\|u\| \le \delta(\varepsilon)$. Entonces, se tiene

$$\begin{aligned} 0 &\le \|L_1(u)-L_2(u)\| \\ &\le \|f(c+u)-f(c)-L_1(u)\|+\|f(c+u)-f(c)-L_2(u)\| \\ &\le 2\varepsilon\,\|u\|. \end{aligned}$$

Por lo tanto, se tiene $0 \le \|L_1(u)-L_2(u)\| \le 2\varepsilon\,\|u\|$ para toda $u \in \mathbf{R}^p$ con $\|u\| \le \delta(\varepsilon)$. Si $L_1 \ne L_2$, existe $z \in \mathbf{R}^p$ con $L_1(z) \ne L_2(z)$, por lo que $z \ne 0$. Ahora sea $z_0=(\delta(\varepsilon)/\|z\|)z$ de manera que $\|z_0\|=\delta(\varepsilon)$ de donde $\|L_1(z_0)-L_2(z_0)\| \le 2\varepsilon\,\|z_0\|$. De modo que $\|L_1(z)-L_2(z)\| \le 2\varepsilon\,\|z\|$ para toda $\varepsilon>0$, de modo que $L_1(z)=L_2(z)$, loque es una contradicción. Por lo tanto, $L_1=L_2$. Q.E.D.

39.4 EJEMPLOS. *(a)* Sea $A \subseteq \mathbf{R}^p$, $y_0 \in \mathbf{R}^q$, y suponga que $f_0 : A \to \mathbf{R}^q$ es la "función constante" definida por $f_0(x)=y_0$ para $x \in A$. Si c es un punto interior de A y $x \in A$, entonces $f_0(x)-f_0(c)=0$. Se deduce que f_0 es diferenciable en c y que la derivada $Df_0(c)=0$, "la función lineal cero" que aplica a todo elemento de $\mathbf{R}^p$ al elemento cero de $\mathbf{R}^q$. *Por lo tanto, la derivada en cualquier punto de una función constante es la función lineal cero.*

(b) Sea $A=\mathbf{R}^p$ y sea $f_1 : A \to \mathbf{R}^q$ una función lineal. Si $c \in A$ y $x \in A$, entonces $f_1(x)-f_1(c)-f_1(x-c)=0$. De aquí que se deduce que f_1 es diferenciable en c y que $Df_1(c)=f_1$. *Por lo tanto, la derivada en cualquier punto de una función lineal es la función lineal misma.*

39.5 LEMA. *Si $f : A \to \mathbf{R}^q$ es diferenciable en $c \in A$, entonces existen números estrictamente positivos δ, K tales que si $\|x-c\| \le \delta$, entonces*

(39.6) $$\|f(x) - f(c)\| \le K\,\|x - c\|.$$

En particular, se deduce que f es continua en $x = c$.

DEMOSTRACION. Por la definición de 39.2 se infiere que existe $\delta > 0$ tal que si $0 < \|x - c\| \le \delta$, entonces (39.2) es válido con $\varepsilon = 1$. Usando la desigualdad del triángulo se tiene

$$\|f(x) - f(c)\| \le \|L(x - c)\| + \|x - c\|$$

para $0 < \|x - c\| \le \delta$. Por el teorema 21.3, existe $B > 0$ tal que $\|L(x - c)\| \le B\,\|x - c\|$ para toda $x \in \mathbf{R}^p$. Por lo tanto, si $0 < \|x - c\| \le \delta$ se obtiene

$$\|f(x) - f(c)\| \le (B + 1)\,\|x - c\|,$$

y esta desigualdad también sigue siendo válida para $x = c$. Q.E.D.

A continuación se demuestra que la existencia de la derivada en un punto implica la existencia de todas las derivadas parciales en ese punto.

39.6 TEOREMA *Si* $A \subseteq \mathbf{R}^p$, *si* $f: A \to \mathbf{R}^q$ *es diferenciable en un punto* $c \in A$, *y si u es cualquier elemento de* $\mathbf{R}^p$, *entonces la derivada parcial* $D_u f(c)$ *de f en c con respecto a u existe. Más aún*

(39.7) $$D_u f(c) = Df(c)(u).$$

DEMOSTRACION. Puesto que f es diferenciable en c, dada $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que

$$\|f(c + tu) - f(c) - Df(c)(tu)\| \le \varepsilon\,\|tu\|$$

siempre que $\|tu\| \le \delta(\varepsilon)$. Si $u = 0$, entonces fácilmente, se puede ver que la derivada parcial con respecto a 0 es $0 = Df(c)(0)$; por lo que se supone que $u \ne 0$. De modo que si $0 < |t| \le \delta(\varepsilon)/\|u\|$, se tiene

$$\left\|\frac{f(c + tu) - f(c)}{t} - Df(c)(u)\right\| \le \varepsilon\,\|u\|.$$

Esto prueba que $Df(c)(u)$ es la derivada parcial de f en c con respecto a u, como se quería. Q.E.D.

39.7 COROLARIO. *Sea* $A \subseteq \mathbf{R}^p$, *sea* $f: A \to \mathbf{R}$ *y sea c un punto interior de A. Si la derivada* $Df(c)$ *existe, entonces cada una de las derivadas parciales* $D_1 f(c), \ldots, D_p f(c)$ *existen en R y si* $u = (u_1, \ldots, u_p) \in \mathbf{R}^p$, *entonces*

(39.8) $$Df(c)(u) = u_1 D_1 f(c) + \cdots + u_p D_p f(c).$$

DEMOSTRACION. El teorema implica que para cada uno de los vectores $e_1, \ldots, e_p$ las derivadas parciales $D_1f(c), \ldots, D_pf(c)$ existen y son iguales a $Df(c)(e_1), \ldots, Df(c)(e_p)$. Sin embargo, dado que $Df(c)$ es lineal y $u = u_1e_1 + \cdots + u_pe_p$, se deduce que

$$Df(c)(u) = \sum_{j=1}^{p} u_j Df(c)(e_j) = \sum_{j=1}^{p} u_j D_j f(c). \qquad \text{Q.E.D.}$$

OBSERVACIONES. *(a)* El inverso del corolario 39.7 no siempre es válido, ya que las derivadas parciales pueden existir sin que exista la derivada. Por ejemplo, defínase $f: \mathbf{R}^2 \to \mathbf{R}$ como

$$\begin{aligned} f(x, y) &= 0 && \text{para } (x, y) = (0, 0), \\ &= \frac{xy^2}{x^2 + y^2} && \text{para } (x, y) \neq (0, 0). \end{aligned}$$

Queda como ejercicio demostrar que la derivada parcial de f con respecto al vector (a, b) en (0,0) está dada por

(39.9) $$D_{(a,b)}f(0, 0) = \frac{ab^2}{a^2 + b^2}, \qquad (a, b) \neq (0, 0).$$

En particular $D_1f(0, 0) = 0$ y $D_2f(0, 0) = 0$. Si la derivada Df existiese en (0,0) el corolario 39.7 implicaría que

$$D_{(a,b)}f(0, 0) = Df(0, 0)(a, b) = a \cdot 0 + b \cdot 0 = 0,$$

contrario a (39.9).

(b) En seguida se verá que si $A \subseteq \mathbf{R}^p$ y si las derivadas parciales de $f: A \to \mathbf{R}^q$ son continuas en c, entonces $Df(c)$ existe.

39.8 EJEMPLOS. *(a)* Sea $A \subseteq \mathbf{R}$ y sea $f: A \to \mathbf{R}$. Entonces, f es diferenciable en un punto interior c de A en el sentido de la definición 39.2 si y sólo si la derivada ordinaria

$$\lim_{\substack{t \to 0 \\ t \neq 0}} \frac{f(c + t) - f(c)}{t} = f'(c)$$

existe. En este caso la derivada $Df(c)$ es la función lineal de $\mathbf{R}$ a $\mathbf{R}$ definida por

$$u \mapsto f'(c)u.$$

De modo que $Df(c)$ aplica $u \in \mathbf{R}$ en el producto de $f'(c)$ y u. (En la terminología de matrices, la derivada $Df(c)$ es la aplicación lineal representada por la matriz 1×1 cuyo único elemento es $f'(c)$.)

Es costumbre que en vez de utilizar u para designar el número real en el que actúa la función lineal de $Df(c)$ se escriba el simbolo un tanto peculiar dx (en este caso la *"d"* desempeña el papel de un prefijo y no tiene ningun otro significado). Cuando esto se

hace y se usa la notación de Liebniz† para la derivada, la fórmula $Df(c)(u)=f'(c)u$ se transforma en

$$Df(c)(dx)=\frac{df}{dx}(c)\,dx.$$

(b) Sea $A\subseteq \boldsymbol{R}$ y sea $f:A\to \boldsymbol{R}^q$ $(q>1)$. De donde f se puede representar por medio de las "funciones coordenadas"

$$f(x)=(f_1(x),\ldots,f_q(x)),\qquad x\in A.$$

Como ejercicio, el lector deberá probar que f es diferenciable en un punto interior c de A si y sólo si cada una de las funciones de valor real $f_1,\ldots,f_q$ tiene una derivada en c. En este caso, la derivada $Df(c)$ es la función lineal de $\boldsymbol{R}$ a $\boldsymbol{R}^q$ dada por

$$u\mapsto u(f_1'(c),\ldots,f_q'(c)),\qquad u\in \boldsymbol{R}.$$

Por lo tanto, $Df(c)$ aplica un número real u en el producto de u y un vector fijo $f'(c)=(f_1'(c),\ldots,f_q'(c))$. Cuando f se considera una "curva", este vector se llama el "vector tangente" a f en el punto $f(c)$.

(c) Sea $A\subseteq \boldsymbol{R}^p$ $(p>1)$ y sea $f:A\to \boldsymbol{R}$. Entonces, del corolario 39.7 se deduce que si la derivada $Df(c)$ existe en un punto c interior a A, entonces cada una de las derivadas parciales $D_1f(c),\ldots,D_pf(c)$ deben existir y $Df(c)$ es la aplicación lineal de $u=(u_1,\ldots,u_p)\in \boldsymbol{R}^p$ a $\boldsymbol{R}$ dada por

$$Df(c)(u)=u_1D_1f(c)+\cdots+u_pD_pf(c).$$

Aun cuando la simple existencia de estas derivadas parciales no implica la existencia de la derivada, se verá más adelante que su continuidad en c sí garantiza la existencia de la derivada.

Algunas veces en vez de $u=(u_1,\ldots,u_p)$ se escribe $dx=(dx_1,\ldots,dx_p)$ para indicar el punto en $\boldsymbol{R}^p$ en que debe actuar la derivada. Cuando esto y la notación de Leibniz se emplean para las derivadas parciales, la fórmula anterior se convierte en

$$Df(c)\,(dx)=\frac{\partial f}{\partial x_1}(c)\,dx_1+\cdots+\frac{\partial f}{\partial x_p}(c)\,dx_p.$$

(d) Suponga ahora $A\subseteq \boldsymbol{R}^p$ y $f:A\to \boldsymbol{R}^q$ en donde ambos $p>1$, $q>1$. En este caso se puede representar $y=f(x)$ por medio de un sistema

$$\begin{aligned} y_1&=f_1(x_1,\ldots,x_p)\\ &\cdots\cdots\cdots\\ y_q&=f_q(x_1,\ldots,x_p), \end{aligned}$$

† GOTTOFRIED WILHELM LEIBNIZ (1646-1716) es, junto con ISAAC NEWTON (1642-1727), uno de los coinventores del cálculo. Leibniz dedicó la mayor parte de su vida a servir a los duquez de Hanover y fue un genio universal Hizo grandes aportaciones en matemáticas, leyes, filosofía, teología, lingüística e historia.

de q funciones de p argumentos. Si f es diferenciable en un punto $(c_1, \ldots, c_p)$ entonces queda como ejercicio demostrar que cada una de las derivadas parciales $D_j f_i(c)(= f_{i,j}(c))$ debe existir en c. (Una vez más, esta última condición no es suficiente, en general, para la diferenciabilidad de f en c.) Cuando $Df(c)$ existe, es la función lineal que aplica al punto $u = (u_1, \ldots, u_p)$ de $\mathbf{R}^p$ en el punto $w = (w_1, \ldots, w_q)$ de $\mathbf{R}^q$ dado por

$$
(39.10) \qquad \begin{aligned} w_1 &= D_1 f_1(c) u_1 + D_2 f_1(c) u_2 + \cdots + D_p f_1(c) u_p, \\ &\cdots\cdots\cdots\cdots\cdots \\ w_q &= D_1 f_q(c) u_1 + D_2 f_q(c) u_2 + \cdots + D_p f_q(c) u_p. \end{aligned}
$$

La derivada $Df(c)$ es la aplicación lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$ determinada por la matriz de $q \times p$ cuyos elementos son

$$
(39.11) \qquad \begin{bmatrix} D_1 f_1(c) & D_2 f_1(c) & \cdots & D_p f_1(c) \\ D_1 f_2(c) & D_2 f_2(c) & \cdots & D_p f_2(c) \\ \cdots & \cdots & \cdots & \cdots \\ D_1 f_q(c) & D_2 f_q(c) & \cdots & D_p f_q(c) \end{bmatrix}
$$

$$
= \begin{bmatrix} f_{1,1}(c) & f_{1,2}(c) & \cdots & f_{1,p}(c) \\ f_{2,1}(c) & f_{2,2}(c) & \cdots & f_{2,p}(c) \\ \cdots & \cdots & \cdots & \cdots \\ f_{q,1}(c) & f_{q,2}(c) & \cdots & f_{q,p}(c) \end{bmatrix}
$$

Ya se ha aclarado (véase el teorema 21.2) que un arreglo tal de números reales determina una función lineal de $\mathbf{R}^p$ a $\mathbf{R}^q$. A la matriz (39.11) se la llama matriz Jacobiana del sistema (39.9) en el punto c. Cuando $p = q$, el determinante de la matriz (39.11) se llama el **determinante Jacobiano** o simplemente el **Jacobiano** del sistema (39.10) en el punto c. Con frecuencia este determinante Jacobiano† se denota por

$$
\left. \frac{\partial(f_1, f_2, \ldots, f_p)}{\partial(x_1, x_2, \ldots, x_p)} \right|_{x=c} \qquad \text{o} \qquad J_f(c).
$$

Existencia de la derivada

En el teorema 39.6 se demostró que la existencia de la derivada en un punto implica la existencia de todas las derivadas parciales en ese punto. En la aclaración que se hizo después del corolario 39.7 se vio que la simple existencia de las derivadas parciales *no* implica la existencia de la derivada aun cuando $p = 2$, $q = 1$. Se habrá de probar ahora que la *continuidad* de las derivadas parciales en c es suficiente para la existencia de la derivada en c.

†CARL (G.J.) JACOBI (1804-1851) fue profesor en Königsberg y Berlín. Su trabajo más importante se refiere a funciones elípticas, pero también se le conoce por su trabajo en determinantes y dinámica.

39.9 TEOREMA. *Sea $A \subseteq \boldsymbol{R}^p$, sea $f: A \to \boldsymbol{R}^q$, y sea c un punto interior de A. Si las derivadas parciales $D_j f_i$ $(i = 1, \ldots, q, j = 1, \ldots, p)$ existen en una vecindad de c y son continuas en c, entonces f es diferenciable en c. Más aún, Df(c) está representada por la matriz* $q \times p$ (39.11).

DEMOSTRACION. Se tratará en detalle el caso $q = 1$ Si $\varepsilon > 0$, sea $\delta(\varepsilon) > 0$ tal que si $\|y - c\| \leq \delta(\varepsilon)$ y $j = 1, 2, \ldots, p$, entonces

$$(39.12) \qquad |D_j f(y) - D_j f(c)| < \varepsilon.$$

Si $x = (x_1, x_2, \ldots, x_p)$ y $c = (c_1, c_2, \ldots, c_p)$, denótense por medio de $z_1, z_2, \ldots, z_{p-1}$ los puntos

$$z_1 = (c_1, x_2, \ldots, x_p), \qquad z_2 = (c_1, c_2, x_3, \ldots, x_p),$$
$$\ldots, z_{p-1} = (c_1, c_2, \ldots, c_{p-1}, x_p)$$

y sean $z_0 = x$ y $z_p = c$. Si $\|x - c\| \leq \delta(\varepsilon)$, entonces es fácil ver que $\|z_j - c\| \leq \delta(\varepsilon)$ para $j = 0, 1, \ldots, p$. La diferencia $f(x) - f(c)$ se escribe como una suma reducida

$$f(x) - f(c) = \sum_{j=1}^{p} \{f(z_{j-1}) - f(z_j)\}.$$

Si se aplica el teorema del valor medio 27.6 al j-ésimo término de esta suma, se obtiene un punto $\bar{z}_j$, que está en el segmento de línea que une a z_{j-1} y z_j, tal que

$$f(z_{j-1}) - f(z_j) = (x_j - c_j) D_j f(\bar{z}_j).$$

Por lo tanto, se obtiene

$$f(x) - f(c) - \sum_{j=1}^{p} (x_j - c_j) D_j f(c) = \sum_{j=1}^{p} (x_j - c_j)\{D_j f(\bar{z}_j) - D_j f(c)\}.$$

Por la desigualdad (39.12), cada cantidad que aparece entre corchetes en la última fórmula está dominada por ε. Aplicando la desigualdad de Schwarz a esta última suma, se obtiene la estima

$$\left\| f(x) - f(c) - \sum_{j=1}^{p} (x_j - c_j) D_j f(c) \right\| \leq (\varepsilon \sqrt{p}) \|x - c\|,$$

siempre que $\|x - c\| \leq \delta(\varepsilon)$.

Se ha demostrado que f es diferenciable en c y que su derivada $Df(c)$ es la función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ dada por

$$u = (u_1, \ldots, u_p) \mapsto Df(c)(u) = \sum_{j=1}^{p} u_j D_j f(c).$$

En el caso en que f toma valores en $\boldsymbol{R}^q$ con $q > 1$, se aplica el mismo argumento a cada una de las funciones de valor real f_i, $i = 1, 2, \ldots, q$, que in-

tervienen en la representación coordenada de la aplicación f. Se dejarán los detalles de este argumento como ejercicio.

Ejercicios

39.A. Defina $f: \boldsymbol{R}^2 \to \boldsymbol{R}$ por medio de

$$f(x, y) = \frac{x}{y} \quad \text{para } y \neq 0,$$
$$= 0 \quad \text{para } y = 0.$$

Demostrar que las derivadas parciales $D_1f(0, 0)$, $D_2f(0, 0)$ existen y son iguales a 0. Sin embargo, la derivada de f en (0,0) con respecto a un vector $u = (a, b)$ no existe si $ab \neq 0$. Demostrar que f no es continua en (0,0); de hecho, f ni siquiera es acotada en una vecindad de (0,0).

39.B. Defina $g: \boldsymbol{R}^2 \to \boldsymbol{R}$ por medio de

$$g(x, y) = 0 \quad \text{para } xy = 0,$$
$$= 1 \quad \text{para } xy \neq 0.$$

Demostrar que las derivadas parciales $D_1g(0, 0)$, $D_2g(0, 0)$ existen y son iguales a 0. Sin embargo, la derivada de g en (0,0) con respecto a un vector $u = (a, b)$ no existe si $ab \neq 0$. Demostrar que g no es continua en (0,0); sin embargo, g es acotada en una vecindad de (0,0).

39.C. Suponga que $h: \boldsymbol{R}^2 \to \boldsymbol{R}$ está definida como

$$h(x, y) = 0 \quad \text{para } (x, y) = (0, 0),$$
$$= \frac{xy}{x^2 + y^2} \quad \text{para } (x, y) \neq (0, 0).$$

Demostrar que las derivadas parciales $D_1h(0, 0)$, $D_2h(0, 0)$ existen y son iguales a 0. Sin embargo, la derivada de h en (0,0) con respecto a un vector $u = (a, b)$ no existe si $ab \neq 0$. Demostrar que h no es continua en (0,0).

39.D. Suponga que $k: \boldsymbol{R}^2 \to \boldsymbol{R}$ está definida como

$$k(x, y) = 0 \quad \text{para } (x, y) = (0, 0),$$
$$= \frac{xy^2}{x^2 + y^4} \quad \text{para } (x, y) \neq (0, 0).$$

Demostrar que la derivada parcial de k en (0,0) con respecto a cualquier vector $u = (a, b)$ existe y que

$$D_u k(0, 0) = \frac{b^2}{a} \quad \text{si } a \neq 0.$$

Demostrar que k no es continua y por lo tanto no es diferenciable en (0,0).

39.E. Defina $f: \boldsymbol{R}^2 \to \boldsymbol{R}$ como

$$f(x, y) = 0 \quad \text{para } (x, y) = (0, 0),$$
$$= \frac{xy^2}{x^2 + y^2} \quad \text{para } (x, y) \neq (0, 0).$$

Demostrar que existe la derivada parcial de f en (0,0) con respecto a cualquier vector $u = (a, b)$ y que

$$D_u f(0, 0) = \frac{ab^2}{a^2 + b^2} \qquad \text{si } (a, b) \neq (0, 0).$$

Demostrar que f es continua pero no diferenciable en (0,0).

39.F. Supóngase que $F: \mathbf{R}^2 \to \mathbf{R}$ está definida como

$$\begin{aligned} F(x, y) &= x^2 + y^2 && \text{si ambos, } x, y \text{ son racionales,} \\ &= 0 && \text{en los otros casos.} \end{aligned}$$

Demostrar que F es continua solamente en el punto (0,0) y que es diferenciable allí.

39.G. Defina $G: \mathbf{R}^2 \to \mathbf{R}$ como

$$\begin{aligned} G(x, y) &= (x^2 + y^2) \operatorname{sen} 1/(x^2 + y^2) && \text{para } (x, y) \neq (0, 0), \\ &= 0 && \text{para } (x, y) = (0, 0). \end{aligned}$$

Demostrar que G es diferenciable en todo punto de $\mathbf{R}^2$ pero las derivadas parciales D_1G, D_2G no son acotadas (y por lo tanto no continuas) en una vecindad de (0,0).

39.H. Defina $H: \mathbf{R}^2 \to \mathbf{R}^2$ como

$$\begin{aligned} H(x, y) &= \left(x^2 + x^2 \operatorname{sen} \frac{1}{x}, y\right) && \text{para } x \neq 0, \\ &= (0, y) && \text{para } x = 0. \end{aligned}$$

Demostrar que D_1H existe en todo punto y que D_2H existe y es continua en una vecindad de (0,0). Demostrar que H es diferenciable en (0,0).

39.I. Sea $A \subseteq \mathbf{R}^p$, sea $f: A \to \mathbf{R}^q$ diferenciable en un punto c interior de A y sea $v \in \mathbf{R}^q$. Si se define $g: A \to \mathbf{R}$ como $g(x) = f(x) \cdot v$ para toda $x \in A$, demostrar que g es diferenciable en c y que

$$Dg(c)(u) = (Df(c)(u)) \cdot v \qquad \text{para } u \in \mathbf{R}^p.$$

39.J. Sea c un punto interior de $A \subseteq \mathbf{R}^p$ y sea $f: A \to \mathbf{R}$.

(a) Si f es diferenciable en c, demostrar que existe un vector único $v_c \in \mathbf{R}^p$ tal que

$$D_u f(c) = Df(c)(u) = v_c \cdot u \qquad \text{para toda } u \in \mathbf{R}^p.$$

Al vector v_c se le llama el **gradiente de** f en c y se denota por medio de $\nabla_c f$, o por grad $f(c)$. Demostrar que

$$\nabla_c f = (D_1 f(c), \ldots, D_p f(c)).$$

(b) Usar la desigualdad de Schwarz para demostrar que si $u \in \mathbf{R}^p$ y $\|u\| = 1$, entonces la función $u \mapsto D_u f(c)$ tiene un valor máximo cuando u es un múltiplo positivo de $\nabla_c f$. De modo que la dirección en que es máxima la derivada direccional de f en c es la del gradiente de f en c.

39.K. Sea c un punto interior de $A \subseteq \mathbf{R}^p$, sean $f, g: A \to \mathbf{R}$ diferenciables en c y sea $\alpha \in \mathbf{R}$. Demostrar que

$$\nabla_c(\alpha f) = \alpha \nabla_c f, \qquad \nabla_c(f+g) = \nabla_c f + \nabla_c g,$$
$$\nabla_c(fg) = f(c)\,\nabla_c g + g(c)\,\nabla_c f.$$

39.L. Encontrar los gradientes de las siguientes funciones en un punto arbitrario en $\mathbf{R}^3$.

(a) $f_1(x, y, z) = x^2 + y^2 + z^2$;
(b) $f_2(x, y, z) = x^2 - yz + z^2$;
(c) $f_3(x, y, z) = xyz$.

39.M. Encontrar las derivadas direccionales de cada una de las funciones en 39.L en el punto (0,1,2,) en la dirección hacia el punto (0,2,3).

39.N. Sea $A \subseteq \mathbf{R}^2$ y suponga que la función $f: A \to \mathbf{R}$ representa una superficie S_f en $\mathbf{R}^3$ explícitamente como su gráfica:

$$S_f = \{(x, y, f(x, y)) : (x, y) \in A\}.$$

Si f es diferenciable en un punto interior (x_0, y_0) de A, entonces el **plano tangente** a S_f en el punto $(x_0, y_0, f(x_0, y_0))$ está dado por la gráfica de la aplicación afín $A_{(x_0,y_0)}: \mathbf{R}^2 \to \mathbf{R}$ definida como

$$A_{(x_0,y_0)}(x, y) = f(x_0, y_0) + Df(x_0, y_0)(x - x_0, y - y_0).$$

Demostrar que el plano tangente a S_f en este punto es

$$\{(x, y, z) \in \mathbf{R}^3 : z = f(x_0, y_0) + D_1 f(x_0, y_0)(x - x_0) + D_2 f(x_0, y_0)(y - y_0)\}.$$

39.O. Encontrar los planos tangentes a las superficies en $\mathbf{R}^3$ representadas como gráficas de las siguientes funciones de los puntos que se especifican. Hacer un dibujo.

(a) $f_1(x, y) = x^2 + y^2$ en $(0, 0)$ y en $(1, 2)$.
(b) $f_2(x, y) = xy$ en $(0, 0)$ y en $(1, 2)$.
(c) $f_3(x, y) = (4 - (x^2 + y^2))^{1/2}$ en $(0, 0)$ y en $(1, 1)$.

39.P. Sea $J \subseteq \mathbf{R}$ un intervalo y suponga que $g: J \to \mathbf{R}^3$ representa a una curva C_g en $\mathbf{R}^3$ paramétricamente:

$$C_g = \{(g_1(t), g_2(t), g_3(t)) : t \in J\}.$$

Si g es diferenciable en un punto interior t_0 de J, entonces el **espacio tangente** a C_g en el punto $g(t_0) = (g_1(t_0), g_2(t_0), g_3(t_0)) \in \mathbf{R}^3$ está dado paramétricamente por la aplicación afín $A_{t_0}: \mathbf{R} \to \mathbf{R}^3$ definida por

$$A_{t_0}(t) = g(t_0) + Dg(t_0)(t - t_0).$$

Demostrar que el espacio tangente a C_g en este punto es

$$\{(x, y, z) \in \mathbf{R}^3 : x = g_1(t_0) + g_1'(t_0)(t - t_0),$$
$$y = g_2(t_0) + g_2'(t_0)(t - t_0), \qquad z = g_3(t_0) + g_3'(t_0)(t - t_0)\}.$$

Si no todos $g_1'(t_0)$, $g_2'(t_0)$, $g_3'(t_0)$ son cero, entonces este espacio tangente es una recta en $\mathbf{R}^3$ y se le llama **recta tangente.**

39.Q. Encontrar ecuaciones paramétricas para las rectas tangentes a las siguien-

tes curvas en $\boldsymbol{R}^3$ en los puntos que se especifican:

(a) $g : t \mapsto (x, y, z) = (t, t^2, t^3)$

en los puntos correspondientes a $t = 0$ y $t = 1$.

(b) $g : t \mapsto (x, y, z) = (t - 1, t^2, 2)$

en los puntos correspondientes a $t = 0$ y $t = 1$.

(c) $g : t \mapsto (x, y, z) = (2 \cos t, 2 \operatorname{sen} t, t)$

en los puntos correspondientes a $t = \pi/2$ y $t = \pi$.

39.R. Sea $A \subseteq \boldsymbol{R}^2$ y suponga que $: A \to \boldsymbol{R}^3$ representa una superficie S_h en $\boldsymbol{R}^3$ paramétricamente:

$$S_h = \{(h_1(s, t), h_2(s, t), h_3(s, t)) : (s, t) \in A\}.$$

Si h es diferenciable en un punto interior (s_0, t_0) de A, entonces el **espacio tangente** a S_h en el punto $h(s_0, t_0) = (h_1(s_0, t_0), h_2(s_0, t_0), h_3(s_0, t_0)) \in \boldsymbol{R}^3$ está dado paramétricamente por la aplicación afín $A_{(s_0, t_0)} : \boldsymbol{R} \to \boldsymbol{R}^3$ definida por

$$A_{(s_0, t_0)}(s, t) = h(s_0, t_0) + Dh(s_0, t_0)(s - s_0, t - t_0).$$

Demostrar que el espacio tangente a S_h en este punto es

$$\begin{aligned}\{(x, y, z) \in \boldsymbol{R}^3 : x &= h_1(s_0, t_0) + D_1h_1(s_0, t_0)(s - s_0) + D_2h_1(s_0, t_0)(t - t_0),\\ y &= h_2(s_0, t_0) + D_1h_2(s_0, t_0)(s - s_1) + D_2h_2(s_0, t_0)(t - t_0),\\ z &= h_3(s_0, t_0) + D_1h_3(s_0, t_0)(s - s_0) + D_2h_3(s_0, t_0)(t - t_0)\}.\end{aligned}$$

Si los vectores $(D_1h_1(s_0, t_0), D_1h_2(s_0, t_0), D_1h_3(s_0, t_0))$ y $(D_2h_1(s_0, t_0), D_2h_2(s_0, t_0), D_2h_3(s_0, t_0))$ en $\boldsymbol{R}^3$ no son múltiplos entre ellos, entonces este espacio tangente es un plano en $\boldsymbol{R}^3$ y recibe el nombre de **plano tangente.**

39.S. Encontrar ecuaciones paramétricas para los planos tangentes a las siguientes superficies en $\boldsymbol{R}^3$ en los puntos que se especifican.

(a) $h : (s, t) \mapsto (x, y, z) = (s, t, s^2 + t^2)$ en los puntos correspondientes a $(s, t) = (0, 0)$ y $(1, 1)$.

(b) $h : (s, t) \mapsto (x, y, z) = (s + t, s - t, s^2 - t^2)$ en los puntos correspondietes a $(s, t) = (0, 0)$ y $(1, 2)$.

(c) $h : (s, t) \mapsto (x, y, z) = (s \cos t, s \operatorname{sen} t, t)$ en los puntos correspondientes a $(s, t) = (1, 0)$ y $(2, \pi/2)$.

(d) $h : (s, t) \mapsto (x, y, z) = (\cos s \operatorname{sen} t, \operatorname{sen} s \operatorname{sen} t,)$ en los puntos correspondientes a $(s, t) = (0, 0)$, $(0, \pi/2)$ y $(\pi/4, \pi/4)$.

39.T. Si $A \subseteq \boldsymbol{R}^p$ y $f : A \to \boldsymbol{R}$ es tal que las derivadas parciales $D_1f, \ldots, D_pf$ existen y son acotadas en alguna vecindad de $c \in A$, entonces f es continua en c. (*Sugerencia:* usar el mismo argumento de la demostración del teorema 39.9.)

39.U. Defina a f en una vecindad de un punto c en $\in \boldsymbol{R}^2$ con valores en $\boldsymbol{R}$. Suponga que D_1f existe y es continua en una vecindad de c y que D_2f existe en c. Demostrar que f es diferenciable en c.

39.V. Sea $A \subseteq \boldsymbol{R}^p$ y suponga que $f : A \to \boldsymbol{R}^q$ y $g : A \to \boldsymbol{R}^r$ están dadas. Si $F : A \to \boldsymbol{R}^q \times \boldsymbol{R}^r = \boldsymbol{R}^{q+r}$ está definida como $F(x) = (f(x), g(x))$ para $x \in A$, demostrar que F es diferenciable en un punto interior $c \in A$ si y sólo si f y g son diferenciables en c. En este caso se tiene

$$DF(c)(u) = \big(Df(c)(u), Dg(c)(u)\big) \quad \text{para } u \in \boldsymbol{R}^p.$$

39.W. Sean $A \subseteq \boldsymbol{R}^p$ y $B \subseteq \boldsymbol{R}^q$ y suponga que $G: A \times B \to \boldsymbol{R}^r$ es diferenciable en un punto (a, b) en $A \times B$. Se definen $g_1: A \to \boldsymbol{R}^r$ y $g_2: B \to \boldsymbol{R}^r$ como las "aplicaciones parciales" en (a, b) dadas por

$$g_1(x) = G(x, b), \qquad g_2(y) = G(a, y)$$

para toda $x \in A$, $y \in B$. Demostrar que g_1 y g_2 son diferenciables en a y b, respectivamente, y que

$$Dg_1(a)(u) = DG(a, b)(u, 0), \qquad Dg_2(b)(v) = DG(a, b)(0, v),$$

para toda $u \in \boldsymbol{R}^p$, $v \in \boldsymbol{R}^q$. Más aún, se tiene

$$DG(a, b)(u, v) = Dg_1(a)(u) + Dg_2(b)(v).$$

(Algunas veces $Dg_1(a) \in \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^r)$ y $Dg_2(b) \in \mathscr{L}(\boldsymbol{R}^q, \boldsymbol{R}^r)$ reciben el nombre de "derivadas parciales del bloque" de G en (a, b) y se denotan por medio de $D_{(1)}G(a, b)$ y $D_{(2)}G(a, b)$.]

Sección 40 La regla de la cadena y teoremas del valor medio

Primero se demostrarán las relaciones algebraicas básicas relacionadas con la derivada. Estas propiedades, que son las mismas que para funciones de valor real de una variable, se usarán con frecuencia en lo sucesivo.

40.1. TEOREMA. *Sea* $A \subseteq \boldsymbol{R}^p$ *y suponga que* c *es un punto interior de* A.

(a) Si f *y* g *están definidas de* A *a* $\boldsymbol{R}^q$ *y son diferenciables en* c *y si* $\beta \in \boldsymbol{R}$, *entonces la función* $h = \alpha f + \beta g$ *es diferenciable en* c *y*

$$Dh(c) = \alpha Df(c) + \beta Dg(c).$$

(b) Si $\varphi: A \to \boldsymbol{R}$ *y* $f: A \to \boldsymbol{R}^q$ *son diferenciables en* c, *entonces la función producto* $k = \varphi f: A \to \boldsymbol{R}^q$ *es diferenciable en* c *y*

$$Dk(c)(u) = \{D\varphi(c)(u)\}f(c) + \varphi(c)\{Df(c)(u)\} \qquad \textit{para } u \in \boldsymbol{R}^p.$$

DEMOSTRACION. *(a)* Si $\varepsilon > 0$, entonces existen $\delta_1(\varepsilon) > 0$ y $\delta_2(\varepsilon) > 0$ tales que si $\|x - c\| \le \inf\{\delta_1(\varepsilon), \delta_2(\varepsilon)\}$, entonces

$$\|f(x) - f(c) - Df(c)(x - c)\| \le \varepsilon \|x - c\|,$$
$$\|g(x) - g(c) - Dg(c)(x - c)\| \le \varepsilon \|x - c\|.$$

De modo que si $\|x - c\| \le \inf\{\delta_1(\varepsilon), \delta_2(\varepsilon)\}$, entonces

$$\|h(x) - h(c) - \{\alpha Df(c)(x - c) + \beta Dg(c)(x - c)\}\| \le (|\alpha| + |\beta|)\varepsilon \|x - c\|.$$

Dado que $\alpha Df(c)+\beta Dg(c)$ es una función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$, se deduce que es diferenciable en c y que $Dh(c)=\alpha Df(c)+\beta Dg(c)$.

(b) Mediante un cálculo sencillo se prueba que

$$\begin{aligned} k(x)-k(c)-\{D\varphi(c)(x-c)f(c)+\varphi(c)Df(c)(x-c)\} \\ =\{\varphi(x)-\varphi(c)-D\varphi(c)(x-c)\}f(x) \\ +D\varphi(c)(x-c)\{f(x)-f(c)\}+\varphi(c)\{f(x)-f(c)-Df(c)(x-c)\}. \end{aligned}$$

Dado que $Df(c)$ existe, del lema 39.5 se deduce que f es continua en c; por lo tanto, existe una constante M tal que $\|f(x)\|<M$ para $\|x-c\|\le\delta$. De aquí se puede ver que todos los términos del lado derecho de la última ecuación se pueden hacer arbitrariamente pequeños escogiendo a $\|x-c\|$ lo suficientemente pequeño. Esto prueba *(b)*. **Q.E.D.**

El siguiente resultado, que es muy importante, asegura que la derivada de la composición de dos funciones diferenciables es la composición de sus derivadas.

40.2 REGLA DE LA CADENA. *Suponga que f tiene dominio $A\subseteq\boldsymbol{R}^p$ y rango en $\boldsymbol{R}^q$, y suponga que g tiene dominio $B\subseteq\boldsymbol{R}^q$ y rango en $\boldsymbol{R}^r$. Sea f diferenciable en c y sea g diferenciable en $b=f(c)$. Entonces, la composición $h=g\circ f$ es diferenciable en c y*

$$Dh(c)=Dg(b)\circ Df(c). \tag{40.1}$$

Alternativamente se escribe

$$D(g\circ f)(c)=Dg(f(c))\circ Df(c). \tag{40.2}$$

DEMOSTRACION. La hipótesis implica que c es un punto interior del dominio de $h=g\circ f$. (¿por qué?) Sea $\varepsilon>0$ y sean $\delta(\varepsilon,f)$ y $\delta(\varepsilon,g)$ como en la definición 39.2. Del lema 39.5 se deduce que existen números estrictamente positivos γ, K tales que si $\|x-c\|\le\gamma$, entonces $f(x)\in B$ y

$$\|f(x)-f(c)\|\le K\,\|x-c\|. \tag{40.3}$$

Para simplificar, se escribe $L_f=Df(c)$ y $L_g=Dg(b)$. Por el teorema 21.3 existe una constante M tal que

$$\|L_g(u)\|\le M\,\|u\|, \quad \text{para } u\in\boldsymbol{R}^q. \tag{40.4}$$

Si $\|x-c\|\le\inf\{\gamma,(1/K)\,\delta(\varepsilon,g)\}$, entonces (40.3) implica que $\|f(x)-f(c)\|\le\delta(\varepsilon,g)$, lo que significa que

$$\|g(f(x))-g(f(c))-L_g(f(x)-f(c))\|\le\varepsilon\,\|f(x)-f(c)\|\le\varepsilon K\,\|x-c\|. \tag{40.5}$$

Si además se pide que $\|x-c\| \leq \delta(\varepsilon, f)$, entonces por (40.4) se infiere que

$$\|L_g\{f(x)-f(c)-L_f(x-c)\}\| \leq \varepsilon M \|x-c\|.$$

Si se combina esta última relación con (40.5) se deduce que si $\delta_1 = \{\gamma, (1/K)\,\delta(\varepsilon, g), \delta(\varepsilon, f)\}$ $x \in A$ y $\|x-c\| \leq \delta_1$, entonces

$$\|g(f(x))-g(f(c))-L_g(L_f(x-c))\| \leq \varepsilon(K+M)\|x-c\|,$$

lo cual significa que

$$\|g \circ f(x)-g \circ f(c)-L_g \circ L_f(x-c)\| \leq \varepsilon(K+M)\|x-c\|.$$

Se concluye que $Dh(c) = L_g \circ L_f$. Q.E.D.

Siguiendo con la notación de la demostración del teorema, $L_f = Df(c)$ una función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$ y $L_g = Dg(b)$ es una función lineal de $\boldsymbol{R}^q$ en $\boldsymbol{R}^r$. La composición $L_g \circ L_f$ es una función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}^r$, como se pide, ya que $h = g \circ f$ es una función definida en una parte de $\boldsymbol{R}^p$ con valores en $\boldsymbol{R}^r$. Se verán ahora algunos ejemplos de este resultado.

40.3 EJEMPLOS. *(a)* Sea $p = q = r = 1$; entonces la derivada $Df(c)$ es una función lineal que manda al número real u hacia $f'(c)u$, y análogamente para $Dg(b)$. Se infiere que la derivada de $g \circ f$ manda al número real u hacia $g'(b)f'(c)u$.

(b) Sea $p > 1$, $q = r = 1$. De acuerdo con el ejemplo 39.8(*c*), la derivada de f en c manda al punto $w = (w_1, \ldots, w_p)$ de $\boldsymbol{R}^p$ hacia el número real

$$D_1 f(c) w_1 + \cdots + D_p f(c) w_p$$

y entonces la derivada de $g \circ f$ en c manda a este punto de $\boldsymbol{R}^p$ hacia el número real

$$g'(b)[D_1 f(c) w_1 + \cdots + D_p f(c) w_p].$$

(c) Sea $q > 1$, $p = r = 1$. De acuerdo con los ejemplos 39.8(*b*), el punto(*c*) la derivada $Df(c)$ manda al número real u hacia el punto

$$Df(c)(u) = uf'(c) = (f_1'(c)u, \ldots, f_q'(c)u) \qquad \text{en } \boldsymbol{R}^q,$$

y la derivada $Dg(b)$ manda al punto $w = (w_1, \ldots, w_q)$ en $\boldsymbol{R}^q$ hacia el número real

$$D_1 g(b) w_1 + \cdots + D_q g(b) w_q.$$

Se infiere que la derivada de $h = g \circ f$ lleva al número real u hacia el número real

$$Dh(c)u = \{D_1 g(b) f_1'(c) + \cdots + D_q g(b) f_q'(c)\}u = u\{Dg(b)(f'(c))\}.$$

La cantidad entre corchetes, que es $h'(c) = (g \circ f)'(c)$ algunas veces se denota de manera menos precisa mediante la notación

$$\frac{\partial g}{\partial y_1}\frac{df_1}{dx}+\cdots+\frac{\partial g}{\partial y_q}\frac{df_q}{dx}$$

En este sentido, ha de quedar claro que las derivadas se deben calcular en puntos apropiados.

(d) Se considera el caso en que $p = q = 2$ y $r = 3$. Para simplificar la notación, las coordenadas variables en $\mathbf{R}^p$ se designarán como (x, y), en $\mathbf{R}^q$ como (x, z) y en $\mathbf{R}^r$ como (r, s, t). Entonces una función f de $\mathbf{R}^p$ a $\mathbf{R}^q$ se puede expresar en la forma

$$w = W(x, y), \qquad z = Z(x, y)$$

y una función g de $\mathbf{R}^q$ a $\mathbf{R}^r$ se puede expresar en la forma

$$r = R(w, z), \qquad s = S(w, z), \qquad t = T(w, z).$$

La derivada $Df(c)$ manda a (ξ, η) hacia (ω, ζ) de acuerdo con las fórmulas

$$\text{(40.6)}\qquad \begin{aligned}\omega &= W_x(c)\xi + W_y(c)\eta,\\ \zeta &= Z_x(c)\xi + Z_y(c)\eta.\end{aligned}$$

Aquí se escribe W_x para denotar $D_1W = D_xW$, etcétera. Asimismo, la derivada $Dg(b)$ manda a (ω, ζ) hacia (ρ, σ, τ) de acuerdo con las relaciones

$$\text{(40.7)}\qquad \begin{aligned}\rho &= R_w(b)\omega + R_z(b)\zeta,\\ \sigma &= S_w(b)\omega + S_z(b)\zeta,\\ \tau &= T_w(b)\omega + T_z(b)\zeta.\end{aligned}$$

Un calculo de rutina prueba que la derivada de $g \circ f$ manda a (ξ, η) hacia (ρ, σ, τ) por medio de

$$\text{(40.8)}\qquad \begin{aligned}\rho &= \{R_w(b)W_x(c) + R_z(b)Z_x(c)\}\xi + \{R_w(b)W_y(c) + R_z(b)Z_y(c)\}\eta,\\ \sigma &= \{S_w(b)W_x(c) + S_z(b)Z_x(c)\}\xi + \{S_w(b)W_y(c) + S_z(b)Z_y(c)\}\eta,\\ \tau &= \{T_w(b)W_x(c) + T_z(b)Z_x(c)\}\xi + \{T_w(b)W_y(c) + T_z(b)Z_y(c)\}\eta.\end{aligned}$$

Una notación más clásica consistiría en escribir dx, dy en vez de ξ, η; dw, dz en vez de ω, ζ; y dr, ds, dt en vez de ρ, σ, τ. Si se designan los valores de la derivada parcial W_x en el punto c por medio de $\partial w/\partial x$, etcétera, entonces (40.6) se transforma en

$$dw = \frac{\partial w}{\partial x}dx + \frac{\partial w}{\partial y}dy,$$

$$dz = \frac{\partial z}{\partial x}dx + \frac{\partial z}{\partial y}dy;$$

análogamente, (40.7) se transforma en

$$dr = \frac{\partial r}{\partial w}dw + \frac{\partial r}{\partial z}dz,$$

$$ds = \frac{\partial s}{\partial w}\, dw + \frac{\partial s}{\partial z}\, dz,$$

$$dt = \frac{\partial t}{\partial w}\, dw + \frac{\partial t}{\partial z}\, dz;$$

y (40.8) se escribe en la forma

$$dr = \left(\frac{\partial r}{\partial w}\frac{\partial w}{\partial x} + \frac{\partial r}{\partial z}\frac{\partial z}{\partial x}\right) dx + \left(\frac{\partial r}{\partial w}\frac{\partial w}{\partial y} + \frac{\partial r}{\partial z}\frac{\partial z}{\partial y}\right) dy,$$

$$ds = \left(\frac{\partial s}{\partial w}\frac{\partial w}{\partial x} + \frac{\partial s}{\partial z}\frac{\partial z}{\partial x}\right) dx + \left(\frac{\partial s}{\partial w}\frac{\partial w}{\partial y} + \frac{\partial s}{\partial z}\frac{\partial z}{\partial y}\right) dy,$$

$$dt = \left(\frac{\partial t}{\partial w}\frac{\partial w}{\partial x} + \frac{\partial t}{\partial z}\frac{\partial z}{\partial x}\right) dx + \left(\frac{\partial t}{\partial w}\frac{\partial w}{\partial y} + \frac{\partial t}{\partial z}\frac{\partial z}{\partial y}\right) dy.$$

En estos tres últimos conjuntos de fórmulas es importante darse cuenta de que todas las derivadas parciales indicadas se deben evaluar en puntos apropiados. Por lo tanto, los coeficientes de dx, dy, etcétera son números reales.

La ecuación (40.6) se puede expresar en términos de una matriz afirmando que la aplicación $Df(c)$ de (ξ, η) en (ω, ζ) está dada por la matriz de 2×2

(40.9)
$$\begin{bmatrix} W_x(c) & W_y(c) \\ Z_x(c) & Z_y(c) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial w}{\partial x}(c) & \dfrac{\partial w}{\partial y}(c) \\ \dfrac{\partial z}{\partial x}(c) & \dfrac{\partial z}{\partial y}(c) \end{bmatrix}.$$

Análogamente, (40.7) asegura que la aplicación $Dg(b)$ de (ω, ζ) en (ρ, σ, τ) está dada por la matriz de 3×2

(40.10)
$$\begin{bmatrix} R_w(b) & R_z(b) \\ S_w(b) & S_z(b) \\ T_w(b) & T_z(b) \end{bmatrix} = \begin{bmatrix} \dfrac{\partial r}{\partial w}(b) & \dfrac{\partial r}{\partial z}(b) \\ \dfrac{\partial s}{\partial w}(b) & \dfrac{\partial s}{\partial z}(b) \\ \dfrac{\partial t}{\partial w}(b) & \dfrac{\partial t}{\partial z}(b) \end{bmatrix}.$$

Por último, la relación (40.8) asegura que la aplicación $D(g\circ f)(c)$ de (ξ, η) en (ρ, σ, τ) está dada por la matriz de 3×2

$$\begin{bmatrix} R_w(b)W_x(c) + R_z(b)Z_x(c) & R_w(b)W_y(c) + R_z(b)Z_y(c) \\ S_w(b)W_x(c) + S_z(b)Z_x(c) & S_w(b)W_y(c) + S_z(b)Z_y(c) \\ T_w(b)W_x(c) + T_z(b)Z_x(c) & T_w(t)W_y(c) + T_z(b)Z_y(c) \end{bmatrix}$$

que es el producto de la matriz en (40.10) con la matriz en (40.9) en ese orden.

Teorema del valor medio

Se pasará ahora al problema de la obtención de una generalización del teorema del valor medio 27.6 para funciones diferenciables de $\mathbf{R}^p$ a $\mathbf{R}^q$. Se verá que el análogo directo del teorema 27.6 no es válido cuando $q > 1$. Se podría esperar que si f es diferenciable en todo punto de $\mathbf{R}^p$ con valores en $\mathbf{R}^q$, y si a, b pertenecen a $\mathbf{R}^p$, entonces existe un punto c (entre a y b) tal que (40.11)

(40.11) $$f(b) - f(a) = Df(c)(b - a).$$

Esta conclusión no es válida aun cuando $p = 1$ y $q = 2$ como se puede ver con la función f definida de $\mathbf{R}$ a $\mathbf{R}^2$ por medio de la fórmula

$$f(x) = (x - x^2, x - x^3).$$

Aquí $Df(c)$ es la función lineal de $\mathbf{R}$ a $\mathbf{R}^2$ que manda al número real u al elemento

$$Df(c)(u) = ((1 - 2c)u, (1 - 3c^2)u).$$

Ahora bien, $f(0) = (0, 0)$ y $f(1) = (0, 0)$, pero no existe ningún punto c tal que $Df(c)(u) = (0, 0)$ para cualquier u distinta de cero en $\mathbf{R}$. Por lo tanto, la fórmula (40.11) no puede ser válida en general cuando $q > 1$, aún cuando $p = 1$. Sin embargo, para muchas aplicaciones basta considerar el caso en que $q = 1$ y entonces es fácil extender el teorema del valor medio.

40.4 TEOREMA DEL VALOR MEDIO. *Sea f una función definida en un subconjunto abierto Ω de $\mathbf{R}^p$ y con valores en $\mathbf{R}$. Suponga que el conjunto Ω contiene a los puntos a,b y al segmento de línea que los une y que f es diferenciable en todos los puntos de este segmento. Entonces existe un punto c en S tal que*

(40.11) $$f(b) - f(a) = Df(c)(b - a).$$

DEMOSTRACION. Suponga que $\varphi : \mathbf{R} \to \mathbf{R}^p$ está definida por medio de

$$\varphi(t) = (1 - t)a + tb = a + t(b - a),$$

de tal manera que $\varphi(0) = a$, $\varphi(1) = b$, y $\varphi(t) \in S \subseteq \Omega$ para $t \in [0, 1]$. Dado que Ω es abierto y φ es continua, existe un número $\gamma > 0$ tal que φ aplica al intervalo $(-\gamma, 1 + \gamma)$ en Ω. Ahora, sea $F : (-\gamma, 1 + \gamma) \to \mathbf{R}$ definida como

$$F(t) = f \circ \varphi(t) = f((1 - t)a + tb).$$

Por la regla de la cadena véase 40.3(c) y 40.P se deduce que

$$\begin{aligned} F'(t) &= Df((1 - t)a + tb)(\varphi'(t)) \\ &= Df((1 - t)a + tb)(b - c). \end{aligned}$$

Si se aplica el teorema del valor medio 27.6 a F se deduce que existe $t_0 \in (0, 1)$ tal que $F(1) - F(0) = F'(t_0)$. Si $c = \varphi(t_0) \in S$, se obtiene

$$\begin{aligned} f(b) - f(a) &= F(1) - F(0) \\ &= F'(t_0) = Df(c)(b-a). \end{aligned}$$

Q.E.D.

Aun cuando la extensión más natural del teorema del valor medio no es válida cuando el espacio del rango es $\mathbf{R}^q$, $q > 1$, hay algunas extensiones que son accesibles. Una de las más útiles se basa en una desigualdad y no en una igualdad.

40.5 TEOREMA DEL VALOR MEDIO. Sea $\Omega \subseteq \mathbf{R}^p$ un conjunto abierto y sea $f: \Omega \to \mathbf{R}^q$. Suponga que Ω contiene a los puntos a, b y al segmento de línea S que une a estos puntos y que f es diferenciable en todo punto de S. Entonces existe un punto c en S tal que

$$\|f(b) - f(a)\| \le \|Df(c)(b-a)\|. \tag{40.12}$$

DEMOSTRACION. Si $y_0 = f(b) - f(a)$ es el vector cero en $\mathbf{R}^q$, entonces el resultado es trivial. Si $y_0 \neq 0$, sea $y_1 = y_0/\|y_0\|$, use el producto interno en $\mathbf{R}^q$ para definir $H: \Omega \to \mathbf{R}$ como

$$H(x) = f(x) \cdot y_1 \qquad \text{para } x \in \Omega.$$

Es evidente que se tiene

$$H(b) - H(a) = \{f(b) - f(a)\} \cdot y_1 = \|f(b) - f(a)\|$$

y fácilmente se puede ver (cf. ejercicio 40.H) que

$$DH(x)(u) = \{Df(x)(u)\} \cdot y_1$$

para $x \in S$, $u \in \mathbf{R}^p$. Del teorema del valor medio 40.4 se deduce que hay un punto c en S tal que

$$\begin{aligned} H(b) - H(a) &= DH(c)(b-a) \\ &= \{Df(c)(b-a)\} \cdot y_1. \end{aligned}$$

Usando la desigualdad de Schwarz y el hecho de que $\|y_1\| = 1$, se tiene

$$\|f(b) - f(a)\| = \{Df(c)(b-a)\} \cdot y_1 \le \|Df(c)(b-a)\|,$$

que es el resultado que se quería. Q.E.D.

Dado que por lo general no se sabe el valor exacto del punto c a menudo se aplica el teorema usando el siguiente resultado, en cuya afirmación se usa el concepto de la norma de una aplicación lineal L de $\mathbf{R}^p$ a $\mathbf{R}^q$ que se intro-

dujo en el ejercicio 21.L. Sólo es necesario aclarar que $\|L(u)\| \leq M\,\|u\|$ para toda $u \in \boldsymbol{R}^p$, si y sólo si la norma $\|L\|_{pq} \leq M$.

40.6 COROLARIO. Suponga que se cumplen las hipótesis del teorema 40.5 y que existe $M > 0$ tal que $\|Df(x)\|_{pq} \leq M$ para toda $x \in S$. Entónces se tiene

$$\|f(b) - f(a)\| \leq M\,\|b - a\|.$$

DEMOSTRACION. Dado que $\|Df(c)(b-a)\| \leq \|Df(c)\|_{pq}\,\|b-a\|$, y como $c \in S$, se tiene y como

$$\|f(b) - f(a)\| \leq \|Df(c)(b-a)\| \leq \|Df(c)\|_{pq}\,\|b-a\| \leq M\,\|b-a\|.$$

Q.E.D.

Intercambio del orden de diferenciación

Si f es una función con dominio en $\boldsymbol{R}^p$ y rango en R, entonces f puede tener p(primeras) derivadas parciales, las que se designan por medio de

$$D_i f \qquad \text{ó} \qquad \frac{\partial f}{\partial x_i}, \qquad i = 1, 2, \ldots, p.$$

Cada una de las derivadas parciales es una función con dominio en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}$ de modo que cada una de estas p funciones puede tener p derivadas parciales. Siguiendo la notación americana aceptada, se hará referencia a las p^2 funciones (o a aquellas que existan) como las **segundas derivadas parciales** de f y se designarán de la siguiente manera

$$D_{ij} f \qquad \text{ó} \qquad \frac{\partial^2 f}{\partial x_j\, \partial x_i}, \qquad i, j = 1, 2, \ldots, p.$$

Se debe observar que la derivada parcial representada por cualquiera de estos últimos símbolos es la derivada parcial con respecto a x_i de la derivada parcial de f con respecto a x_i. (En otras palabras, primero x_i, y después x_j).

De manera análoga se puede analizar la existencia de las terceras derivadas parciales y de aquellas de mayor orden. En principio, una función de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ puede tener tantas como p^n n-ésimas derivadas parciales. Sin embargo, es muy conveniente que si las derivadas que resultan son continuas entonces el orden de diferenciación no es significativo. Además de que hace que disminuya el número de de derivadas parciales de orden alto (potencialmente distintas), este resultado elimina en gran parte el peligro de la distinción de notación un tanto sutil que se emplea para distintos órdenes de diferenciación.

Basta con tomar en cuenta el intercambio de orden para segundas derivadas. Manteniendo a todas las otras coordenadas constantes, se puede ver que no se pierde generalidad al considerar una función de $\mathbf{R}^2$ a $\mathbf{R}$. Con el objeto de simplificar la notación se denota por medio de (x, y) un punto en $\mathbf{R}^2$ y se habrá de probar que si existen $D_x f, D_y f$, y si $D_{yx} f$ es continua en un punto, entonces la derivada parcial $D_{yx} f$ existe en este punto y es igual a $D_{xy} f$ En el ejercicio 40.U se verá que es posible que tanto $D_{yx} f$ como $D_{xy} f$ existan en un punto y sin embargo no sean iguales.

El camino que se usará en esta demostración es el de probar que estas dos derivadas parciales mixtas en el punto (0,0) son el límite del cociente

$$\frac{f(h, k) - f(h, 0) - f(0, k) + f(0, 0)}{hk},$$

conforme (h, k) se aproxima a (0,0).

40.7 LEMA. *Suponga que f está definida en una vecindad U del origen en $\mathbf{R}^2$ con valores en $\mathbf{R}$, que las derivadas parciales $D_x f$ y $D_{yx} f$ existen en U y que $D_{yx} f$ es continua en $(0,0)$. Si A es la diferencia compuesta*

(40.13) $$A(h, k) = f(h, k) - f(h, 0) - f(0, k) + f(0, 0),$$

entonces se tiene

$$D_{yx} f(0, 0) = \lim_{(h, k) \to (0, 0)} \frac{A(h, k)}{hk}.$$

DEMOSTRACION. Sean $\varepsilon > 0$ y $\delta > 0$ tan pequeños que si $|h| < \delta$ y $|k| < \delta$, entonces el punto (h, k) pertenece a U y

(40.14) $$|D_{yx} f(h, k) - D_{yx} f(0, 0)| < \varepsilon.$$

Si $|k| < \delta$, se define B para $|h| < \delta$ por medio de

$$B(h) = f(h, k) - f(h, 0),$$

de donde se infiere que $A(h, k) = B(h) - B(0)$. Por hipótesis, la derivada parcial $D_x f$ existe en U y por lo tanto B tiene una derivada. Aplicando el teorema del valor medio 27.6 a B existe un número h_0 con $0 < |h_0| < |h|$ tal que

(40.15) $$A(h, k) = B(h) - B(0) = hB'(h_0).$$

(Se observa que el valor de h_0 depende del valor de k, pero esto no ocasionará ningún problema.) Haciendo referencia a la definición de B se tiene

$$B'(h_0) = D_x f(h_0, k) - D_x f(h_0, 0).$$

Aplicando el teorema del valor medio al lado derecho de la última ecuación existe un número k_0 con $0 < |k_0| < |k|$ tal que

(40.16) $$B'(h_0) = k\{D_{yx}f(h_0, k_0)\}.$$

Combinando las ecuaciones (40.15) y (40.16) se concluye que si $0<|h|<\delta$ y $0<|k|<\delta$, entonces

$$\frac{A(h, k)}{hk} = D_{yx}f(h_0, k_0),$$

en donde $0<|h_0|<|h|$, $0<|k_0|<|k|$. Se sigue de la desigualdad (40.14) y la expresión anterior que

$$\left|\frac{A(h, k)}{hk} - D_{yx}f(0, 0)\right| < \varepsilon$$

siempre que $0<|h|<\delta$ y $0<|k|<\delta$. Q.E.D.

Se puede obtener ahora una condición útil (debida a H.A. Schwarz) para la igualdad de las parciales mixtas.

40.8 TEOREMA. *Suponga que f está definida en una vecindad U de un punto (x, y) en $\mathbf{R}^2$ con valores en $\mathbf{R}$. Suponga que las derivadas parciales D_xf, D_yf, y $D_{yx}f$ existen en U y que $D_{xy}f$ es continua en (x, y). Entonces la derivada parcial $D_{yx}f$ existe en $D_{xy}f(x, y) = D_{yx}f(x, y)$.*

DEMOSTRACION. Sin perder generalidad se puede suponer que $(x, y) = (0, 0)$ y es lo que se hará. Si A es la función definida en el lema anterior, entonces se vio que

(40.17) $$D_{yx}f(0, 0) = \lim_{(h,k)\to(0,)} \frac{A(h, k)}{hk},$$

siendo la existencia de este límite parte de la conclusión. Por hipótesis D_yf existe en U, de manera que

(40.18) $$\lim_{k\to 0} \frac{A(h, k)}{hk} = \frac{1}{h}\{D_yf(h, 0) - D_yf(0, 0)\}, \qquad h \neq 0.$$

Si $\varepsilon > 0$, existe un número $\delta(\varepsilon) > 0$ tal que si $0<|h|<\delta(\varepsilon)$ y $0<|k|<\delta(\varepsilon)$, entonces

$$\left|\frac{A(h, k)}{hk} - D_{yx}f(0, 0)\right| < \varepsilon.$$

Tomando el límite en esta desigualdad con respecto a k y usando (40.18), se obtiene

$$\left|\frac{1}{h}\{D_yf(h, 0) - D_yf(0, 0)\} - D_{yx}f(0, 0)\right| \leq \varepsilon,$$

para toda h que satisfaga $0<|h|<\delta(\varepsilon)$. Por lo tanto, $D_{xy}f(0, 0)$ existe y es igual a $D_{yx}f(0, 0)$ Q.E.D.

Derivadas de orden superior

Si f es una función con dominio en $\boldsymbol{R}^p$ y rango en $\boldsymbol{R}$, entonces la derivada $Df(c)$ de f en c es la función lineal de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ tal que

$$\|f(c+z)-f(c)-Df(c)(z)\| \le \varepsilon \|z\|,$$

para z suficientemente pequeña. Esto significa que $Df(c)$ es la función lineal, que más se aproxima a la diferencia $f(c+z)-f(c)$ cuando z es pequeña. Cualquier otra función lineal llegaría a una aproximación menos exacta para z pequeña. A partir de esta propiedad descriptiva se puede ver que si $Df(c)$ existe, entonces necesariamente está dada por la fórmula

$$Df(c)(z) = D_1f(c)z_1 + \cdots + D_pf(c)z_p,$$

en donde $z=(z_1, \ldots, z_p)$ en $\boldsymbol{R}^p$.

A pesar de que las aproximaciones lineales son en particular sencillas y son lo suficientemente exactas para muchos propósitos, algunas veces es deseable obtener un mayor grado de aproximación usando funciones lineales. En tales casos resulta natural dirigirse a funciones cuadráticas, funciones cúbicas, etcétera, para efectuar aproximaciones más precisas. Dado que las funciones que aquí se tratan tienen sus dominios en $\boldsymbol{R}^p$, sería necesario el estudio de funciones multilineales de $\boldsymbol{R}^p$ a $\boldsymbol{R}$ para un análisis cabal de dichas funciones. Aun cuando dicho estudio no es particularmente difícil, significaría salirse mucho del tema dadas las aplicaciones limitadas que se tienen en mente.

Por este motivo, se definirá la **segunda derivada** $D^2f(c)$ de f en c como la función de $\boldsymbol{R}^p \times \boldsymbol{R}^p$ a $\boldsymbol{R}$ tal que si (y, z) pertenece a este producto y $y=(y_1, \ldots, y_p)$ y $z=(z_1, \ldots, z_p)$, entonces

$$D^2f(c)(y, z) = \sum_{i,j=1}^{p} D_{ji}f(c)y_iz_j.$$

Al estudiar la segunda derivada se habrá de suponer en lo subsiguiente que las segundas derivadas parciales de f existen y son continuas en una vecindad de c. De manera análoga se define la **tercera derivada** $D^3f(c)$ de f en c como la función de (y,z,w) en $\boldsymbol{R}^p \times \boldsymbol{R}^p \times \boldsymbol{R}^p$ dada por

$$D^3f(c)(y, z, w) = \sum_{i,j,k=1}^{p} D_{kji}f(c)y_iz_jw_k.$$

Al analizar la tercera derivada se habrá de suponer que todas las terceras derivadas parciales de f existen y son continuas en una vecindad de c.

Hasta aquí, el método de formación de las derivadas de orden superior ya debe estar claro. (Dadas las observaciones anteriores con respecto al intercambio de orden en la diferenciación, si las derivadas parciales mixtas que resultan son continuas, entonces son independientes del orden de diferenciación.)

Otra notación: se escribe

$$\begin{array}{lll} D^2f(c)(w)^2 & \text{para} & D^2f(c)(w, w), \\ D^3f(c)(w)^3 & \text{para} & D^3f(c)(w, w, w), \\ \cdots & & \\ D^nf(c)(w)^n & \text{para} & D^nf(c)(w, w, \ldots, w). \end{array}$$

Si $p = 2$ y si se designa a un elemento de $\mathbf{R}^2$ como (x, y) y $w = (h, k)$, entonces $D^2f(c)(w)^2$ es igual a la expresión

$$D_{xx}f(c)h^2 + 2D_{xy}f(c)hk + D_{yy}f(c)k^2;$$

en forma análoga $D^3f(c)(w)^3$ es igual a

$$D_{xxx}f(c)h^3 + 3D_{xxy}f(c)h^2k + 3D_{xyy}f(c)hk^2 + D_{yyy}f(c)k^3,$$

y $D^nf(c)(w)^n$ es igual a la expresión.

$$D_{x\cdots x}f(c)h^n + \binom{n}{1}D_{x\cdots xy}f(c)h^{n-1}k + \binom{n}{2}D_{x\cdots xyy}f(c)h^{n-2}k^2 + \cdots + D_{y\cdots y}f(c)k^n.$$

Una vez introducida esta notación se habrá de probar una generalización importante del teorema de Taylor para funciones de $\mathbf{R}^p$ a $\mathbf{R}$.

40.9 TEOREMA DE TAYLOR. *Suponga que f es una función con dominio abierto Ω en $\mathbf{R}^p$ y rango en $\mathbf{R}$ y suponga que f tiene derivadas parciales continuas de orden n en una vecindad de todo punto en un segmento de línea S que una a dos puntos a, $b = a + u$ en Ω. Entonces existe un punto c en S tal que*

$$f(a+u) = f(a) + \frac{1}{1!}Df(a)(u) + \frac{1}{2!}D^2f(a)(u)^2 + \cdots + \frac{1}{(n-1)!}D^{n-1}f(a)(u)^{n-1} + \frac{1}{n!}D^nf(c)(u)^n.$$

DEMOSTRACION. Suponga que F esta definida para t de $\mathbf{I}$ a $\mathbf{R}$ por medio de

$$F(t) = f(a + tu).$$

Por el supuesto de que existen las derivadas parciales de f se sigue que

$$\begin{aligned} F'(t) &= Df(a+tu)(u), \\ F''(t) &= D^2f(a+tu)(u)^2, \\ &\cdots \\ F^{(n)}(t) &= D^nf(a+tu)(u)^n. \end{aligned}$$

Si se aplica la versión unidimensional del teorema de Taylor 28.6 a la función F en I, se infiere que existe un número real t_0 en I tal que

$$F(1)=F(0)+\frac{1}{1!}F'(0)+\cdots+\frac{1}{(n-1)!}F^{(n-1)}(0)+\frac{1}{n!}F^{(n)}(t_0).$$

Si se fija $c=a+t_0u$, el resultado se deduce Q.E.D.

Ejercicios

40.A. Si $f(x, y)=x^2+y^2$ y $g(t)=(3t+1,\ 2t-3)$, sea $F(t)=f\circ g(t)$. Calcular $F'(t)$ directamente y usando la regla de la cadena.

40.B. Si $f(x, y)=xy$ y $g(s, t)=(2s+3t, 4s+t)$, sea $F(s, t)=f\circ g(s, t)$. Evaluar D_1F y D_2F directamente y usando la regla de la cadena.

40.C. Si $f(x, y, z)=xyz$ y $g(s, t)=(3s+st, s, t)$, sea $F(s, t)=f\circ g(s, t)$. Evaluar D_1F y D_2F directamente y usando la regla de la cadena.

40.D. Si $f(x, y, z)=xy+yz+zx$ y $g(s, t)=(\cos s, \operatorname{sen} s \cos t, \operatorname{sen} t)$, sea $F(s, t)=f\circ g(s, t)$. Evaluar D_1F y D_2F directamente usando la regla de la cadena.

40.E. Si se hacen rotar los ejes de coordenadas en el plano con un ángulo θ, entonces las nuevas coordenadas u, v del punto están relacionadas a las coordenadas originales x, y de la siguiente manera

$$x=u\cos\theta-v\operatorname{sen}\theta,\qquad y=u\operatorname{sen}\theta+v\cos\theta.$$

Sea $f:\boldsymbol{R}^2\to\boldsymbol{R}$ diferenciable en $\boldsymbol{R}^2$ y sea $F(u, v)=f(x, y)$ para toda x, y. Demostrar que

$$[D_1F(u, v)]^2+[D_2F(u, v)]^2=[D_1f(x, y)]^2+[D_2f(x, y)]^2.$$

40.F. Sea $f:\boldsymbol{R}^2\to\boldsymbol{R}$ diferenciable en $\boldsymbol{R}^2$, suponga que $g:(0,+\infty)\times\boldsymbol{R}\to\boldsymbol{R}$ está definida como $g(r, \theta)=(r\cos\theta, r\operatorname{sen}\theta)$, y sea $F=f\circ g$. Calcular D_1F y D_2F y demostrar que

$$[D_1F(r, \theta)]^2+\frac{1}{r^2}[D_2F(r, \theta)]^2=[D_1f(r\cos\theta, r\operatorname{sen}\theta)]^2+[D_2f(r\cos\theta, r\operatorname{sen}\theta)]^2.$$

40.G. Sea $f:\boldsymbol{R}\to\boldsymbol{R}$ diferenciable en $\boldsymbol{R}$.

(a) Si $F(x, y)=f(xy)$, entonces $xD_1F(x, y)=yD_2F(x, y)$ para todo (x, y).

(b) Si $F(x, y)=f(ax+by)$ en donde $a, b\in\boldsymbol{R}$, entonces $bD_1F(x, y)=aD_2F(x, y)$ para todo (x, y).

(c) Si $F(x, y)=f(x^2+y^2)$, entonces $yD_1F(x, y)=xD_2F(x, y)$ para todo (x, y).

(d) Si $F(x, y)=f(x^2-y^2)$, entonces $yD_1F(x, y)+xD_2F(x, y)=0$ para todo (x, y).

40.H. Sea $A\subseteq\boldsymbol{R}^p$ y sea c un punto interior de A. Suponga que f, g están definidas de A a $\boldsymbol{R}^q$ y son diferenciables en c. Si $h:A\to\boldsymbol{R}$ está definida como $h(x)=f(x)\cdot g(x)$ para toda $x\in A$, demostrar que h es diferenciable en c y que si $u\in\boldsymbol{R}^p$, entonces

$$Dh(c)(u) = (Df(c)(u)) \cdot g(c) + f(c) \cdot (Dg(c)(u)).$$

40.I. Expresar el resultado del ejercicio 40.H en términos de las funciones coordenadas.

40.J. Sea $A \subseteq \boldsymbol{R}$ y sea c un punto interior de A. Suponga que $f : A \to \boldsymbol{R}^p$ es diferenciable en c y es tal que $\|f(x)\| = 1$ para $x \in A$. Demostrar que $f(c) \cdot \nabla_c f = 0$, en donde $\nabla_c f$ designa el gradiente de f en c (véase el ejercicio 39.J). Hacer una interpretación geométrica de esta conclusión.

40.K. Sea $f : \boldsymbol{R}^p \to \boldsymbol{R}$ (absolutamente) **homogénea de grado** k en el sentido de que

$$f(tx) = t^k f(x) \qquad \text{para} \qquad x \in \boldsymbol{R}^p,\ t > 0.$$

(a) Si f es diferenciable en $\boldsymbol{R}^p$, demostrar que satisface la **relación de Euler**†:

$$kf(x) = x_1 D_1 f(x) + \cdots + x_p D_p f(x)$$

para toda $x = (x_1, \ldots, x_p)$ en $\boldsymbol{R}^p$ con $x \neq 0$.

(b) Inversamente sea f tal que satisface la relación de Euler y sea $c \in \boldsymbol{R}^p$, $c \neq 0$. Sea $g(t) = f(tc)$ para $t > 0$ demostrar que $tg'(t) = kg(t)$ para $t > 0$. Usar esto para demostrar que f es homogénea de grado k.

40.L. Sea $A \subseteq \boldsymbol{R}^p$, $f : A \to \boldsymbol{R}^p$, y suponga que la función $g : f(A) \to \boldsymbol{R}^p$ es inversa de f en el sentido de que

$$f \circ g(x) = x, \qquad g \circ f(y) = y$$

para toda $x \in A$ y $y \in f(A)$. Si f es diferenciable en un punto $a \in A$ y si g es diferenciable en $b = f(a)$, demostrar que las funciones lineales $Df(a)$ y $Dg(b)$ son inversas entre si, es decir $Df(a) \circ Dg(b)$ y $Dg(b) \circ Df(a)$ son la identidad en $\boldsymbol{R}^p$.

40.M. Sea $B : \boldsymbol{R}^p \times \boldsymbol{R}^p = \boldsymbol{R}^{2p} \to \boldsymbol{R}^q$ bilineal en el sentido de que

$$B(ax + bx', y) = aB(x, y) + bB(x', y),$$
$$B(x, ay + by') = aB(x, y) + bB(x, y')$$

para todas $a, b \in \boldsymbol{R}$ y todas x, x', y, y' en $\boldsymbol{R}^p$. Se puede demostrar que existe $M > 0$ tal que $\|B(x, y)\| \leq M \|x\| \|y\|$ para todas x, y en $\boldsymbol{R}^p$. Suponiendo esto, demostrar que B es diferenciable en todo punto $(x, y) \in \boldsymbol{R}^p \times \boldsymbol{R}^p = \boldsymbol{R}^{2p}$ y que

$$DB(x, y)(u, v) = B(x, v) + B(u, y)$$

para todo (u, v) en $\boldsymbol{R}^p \times \boldsymbol{R}^p = \boldsymbol{R}^{2p}$.

40.N. Sea $B : \boldsymbol{R}^p \times \boldsymbol{R}^p \to \boldsymbol{R}^q$ bilineal en el mismo sentido del ejercicio anterior y sea $g(x) = B(x, x)$ para toda $x \in \boldsymbol{R}^p$. Demostrar que si x, $u \in \boldsymbol{R}^p$, entonces

(i) $g(tx) = t^2 g(x)$ para toda $t \in \boldsymbol{R}$;
(ii) $Dg(x)(u) = B(x, u) + B(u, x) = Dg(u)(x)$;
(iii) $g(x + u) = g(x) + Dg(x)(u) + g(u)$.

Más aún, si B es simétrica en el sentido de que $B(x, y) = B(y, x)$, entonces

† LEONARD EULER (1707-1783), originario de Basel, estudió con Johann Bernoulli. Residió muchos años en la corte en San Petersburgo, pero su estancia fue interrumpida por veinticinco años en Berlín. No obstante el hecho de haber sido padre de trece h[illegible]s y de haber perdido la vista totalmente, fue capaz de escribir más de ochocientos trabajos y libros y de hacer aportaciones fundamentales a todas las ramas de las matemáticas.

(iv) $Dg(x)(u) = 2B(x, u)$.

40.O. Dar una demostración del ejercicio 40.H usando 40.M.

40.P. Sea $\Omega \subseteq R^p$ abierto y sea $f : \Omega \to R^q$ diferenciable en Ω. Suponga que $I = (a, b)$ es un intervalo abierto en R y sea $g : I \to R^p$ diferenciable en I y tal que $g(I) \subseteq \Omega$. Si $h = f \circ g : I \to R^q$ demostrar que

$$h'(c) = Df(g(c))(g'(c)).$$

40.Q. Sea $\Omega \subseteq R^p$ abierto y sea $f : \Omega \to R^q$. Suponga que Ω contiene a los puntos a, b y al segmento de línea S que une a estos puntos y que es diferenciable en todo punto de S. Demostrar que existe una aplicación lineal $L : R^p \to R^q$ tal que $f(b) - f(a) = L(b - a)$.

40.R. Sea $\Omega \subseteq R^p$ un conjunto conexo abierto y sea $f : \Omega \to R^q$ diferenciable en Ω. Si $Df(x) = 0$ para toda $x \in \Omega$, demostrar que $f(x) = f(y)$ para toda $x, y \in \Omega$. Demostrar que esta conclusión puede no ser válida si Ω no es conexo.

40.S. Sea $J \subseteq R^p$ una celda abierta y suponga que $f : J \to R$ es diferenciable en J. Demostrar que si la derivada parcial $D_1 f(x) = 0$ para toda $x \in J$, entonces f no depende de la primera variable en el sentido de que

$$f(x_1, x_2, \ldots, x_p) = f(x_1', x_2, \ldots, x_p)$$

para cualesquiera dos puntos en J cuyas segunda,..., p-ésima coordenadas sean la misma.

40.T. Demostrar que la conclusión del ejercicio anterior puede no ser válida si no se supone que J es una celda.

40.U. Suponga que $f : R^2 \to R$ está definida como

$$f(x, y) = \frac{xy(x^2 - y^2)}{x^2 + y^2} \quad \text{para } (x, y) \neq (0, 0),$$
$$= 0 \quad \text{para } (x, y) = (0, 0).$$

Demostrar que las segundas derivadas parciales $D_{xy}f$ y $D_{yx}f$ existen en (0,0) pero no son iguales.

40.V. Usar el teorema del valor medio para determinar aproximadamente la distancia del punto (3.2,4.1) al origen. Dar cotas de error para la estima.

40.W. Sea $\Omega \subseteq R^p$ abierto y sea $f : \Omega \to R^q$. Supóngase que Ω contiene a los puntos a, b, al segmento de línea S que los une y que f tiene derivadas parciales continuas en S. Demostrar que

$$f(b) - f(a) = \int_0^1 Df(a + t(b - a))(b - a)\, dt.$$

40.X. Supóngase que f, $g : R \to R$ tienen segundas derivadas continuas en R.

(a) Si $c \in R$ y $u(x, y) = f(x + cy) + g(x - cy)$, demostrar que $u : R^2 \to R$ satisface la "ecuación de onda"

$$c^2 D_{xx}u(x, y) = D_{yy}u(y, y)$$

para todo (x, y).

(b) Si $v(x, y) = f(3x + 2y) + g(x - 2y)$, demostrar que $v : R^2 \to R$ satisface la ecuación

$$4D_{xx}v(x, y) - 4D_{xy}v(x, y) - 3D_{yy}v(x, y) = 0$$

para todo (x, y).

40.Y. Si $f: \mathbf{R}^2 \to \mathbf{R}$ tiene segundas derivadas parciales continuas y si $F(r, \theta) = f(r \cos \theta, r \operatorname{sen} \theta)$ para $r > 0$, $\theta \in \mathbf{R}$, demostrar que

$$D_{xx}f(x, y) + D_{yy}f(x, y) = D_{rr}F(r, \theta) + \frac{1}{r} D_r F(r, \theta) + \frac{1}{r^2} D_{\theta\theta}F(r, \theta)$$
$$= \frac{1}{r} D_r(rD_rF(r, \theta)) + \frac{1}{r^2} D_{\theta\theta}F(r, \theta),$$

en donde $x = r \cos \theta$, $y = r \operatorname{sen} \theta$.

Proyecto

40.α. (Este proyecto es una modificación del clásico **método de Newton** para la ubicación de ráices cuando se conoce una aproximación suficientemente cercana.) Suponga que f está definida y es continua en un conjunto abierto que contiene a la bola cerrada $B_r(x_0) = \{x \in \mathbf{R}^p : \|x - x_0\| \leq r\}$ con valores en $\mathbf{R}^q$. Suponga que f es diferenciable en todo punto de $B_r(x_0)$ y que existe un número C, con $0 < C < 1$, y una aplicación lineal inyectiva $\Gamma: \mathbf{R}^q \to \mathbf{R}^p$ tal que $\|\Gamma \circ f(x_0)\| \leq (1 - C)r$ y tal que

$$\|I - \Gamma \circ Df(x)\|_{pp} \leq C \quad \text{para } x \in B_r(x_0).$$

(a) Sea $g: B_r(x_0) \to \mathbf{R}^p$ la función definida por $g(x) = x - \Gamma \circ f(x)$ para $x \in B_r(x_0)$. Demostrar que g es diferenciable en todo punto de $B_r(x_0)$ y que g es una contracción con constante $C < 1$ (véase 23.4) en $B_r(x_0)$.

(b) Defina $x_1 = g(x_0)$ y $x_{n+1} = g(x_n)$ para $n \in \mathbf{N}$. Demostrar que $\|x_{n+1} - x_n\| \leq C^n \|x_1 - x_0\|$, de donde se infiere que $\|x_{n+1} - x_m\| \leq C^m r$ para $n \geq m \geq 0$. Por lo tanto, $\|x_k - x_0\| < r$ para $k = 0, 1, 2, \ldots$.

(c) Demostrar que (x_k) es una sucesión de Cauchy y por lo tanto converge a un elemento $\bar{x} \in B_r(x_0)$, que es tal que $g(\bar{x}) = \bar{x}$. Más aún, se tiene la estima $\|x_k - \bar{x}\| \leq \quad C^k r$.

(d) Demostrar que $f(\bar{x}) = 0$ y $\bar{x}$ es el único elemento en $B_r(x_0)$ en donde f desaparece

Sección 41 Teoremas sobre transformaciones y funciones implícitas

Sea $\boldsymbol{\Omega}$ un conjunto abierto en $\mathbf{R}^p$ y sea f una función con dominio $\boldsymbol{\Omega}$ y rango en $\mathbf{R}^q$; a reserva de que se especifique lo contrario, no se supondrá que $p = q$. Se probará que, con supuestos que se establecerán, el "caracter local" de la transformación f en un punto $c \in \boldsymbol{\Omega}$ está indicado por la transformación lineal $Df(c)$. De manera un poco más precisa:

(i) si $p \leq q$ y $Df(c)$ es inyectiva (= uno a uno), entonces f es inyectiva en vecindades pequeñas de c;

(ii) si $p \geq q$ y $Df(c)$ es suprayectiva (= aplica a $\boldsymbol{R}^p$ sobre $\boldsymbol{R}^q$), entonces la imagen bajo f de una vecindad pequeña de c es una vecindad de $f(c)$ y

(iii) si $p = q$ y $Df(c)$ es biyectiva (= uno a uno y sobre = invertible), entonces f aplica una vecindad U de c en una forma uno a uno sobre una vecindad V de $f(c)$. En el caso *(iii)* hay una función definida en V que es inversa a la restricción de f a U.

Como una consecuencia de estos teoremas sobre transformaciones, se obtendrá el teorema de la función implícita que es uno de los teoremas fundamentales en análisis y geometría. También se ofrecerá un útil teorema de parametrización y el importante teorema del rango.

La clase $C^1(\Omega)$

Tan sólo la existencia de la derivada no es suficiente para los propósitos que aquí se tienen, también es necesaria la continuidad de la derivada. Recuerde que si $f:\Omega \to \boldsymbol{R}^q$ es diferenciable en todo punto de $\Omega \subseteq \boldsymbol{R}^p$, entonces la función $x \mapsto Df(x)$ es una aplicación de Ω hacia la colección $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ de todas las transformaciones lineales de $\boldsymbol{R}^p$ a $\boldsymbol{R}^q$. En la sección 21 se hizo la observación de que este conjunto $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ es un espacio vectorial y en el ejercicio 21.L, que este espacio es un espacio normado según la norma.

$$\|L\|_{pq} = \sup \{\|L(x)\| : x \in \boldsymbol{R}^p, \|x\| \leq 1\}. \tag{41.1}$$

41.1 DEFINICION. Si Ω es abierto en $\boldsymbol{R}^p$ y $f:\Omega \to \boldsymbol{R}^q$, se dice que f pertenece a la clase $C^1(\Omega)$ si la derivada $Df(x)$ existe para toda $x \in \Omega$ y la transformación $x \mapsto Df(x)$ de Ω en $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ es continua bajo la norma (41.1)

Recuerde del ejemplo 39.8(*d*), que para cada $x \in \Omega$, la derivada $Df(x)$ se puede representar por medio de la matriz jacobiana de $q \times p\,[D_jf_i(x)]$. Por lo que $Df(x) - Df(y)$ se representa por la matriz de $q \times p$

$$[D_jf_i(x) - D_jf_i(y)].$$

Ahora, de la desigualdad (21.5) se sigue que

$$\|Df(x) - Df(y)\|_{pq} \leq \left\{\sum_{i=1}^{q} \sum_{j=1}^{p} |D_jf_i(x) - D_jf_i(y)|^2\right\}^{1/2}$$

De donde la continuidad de cada una de las derivadas parciales D_jf_i en Ω implica la continuidad de $x \mapsto Df(x)$. Se le deja al lector demostrar que lo inverso también es válido. Por lo tanto, se tiene el siguiente resultado.

41.2 TEOREMA. *Si $\Omega \subseteq \boldsymbol{R}^p$ es abierto y $f:\Omega \to \boldsymbol{R}^q$ es diferenciable en todo punto de Ω, entonces f pertenece a la clase $C^1(\Omega)$ si y sólo si las derivadas parciales D_jf_i, $i = 1, \ldots, q$, $j = 1, \ldots, p$, de f son continuas en Ω.*

El siguiente lema que es una variante del teorema del valor medio será necesario.

41.3 LEMA. *Sea $\Omega \subseteq \boldsymbol{R}^p$ un conjunto abierto y sea $f:\Omega \to \boldsymbol{R}^q$ diferenciable en Ω. Suponga que Ω contiene a los puntos a, b y al segmento de línea S que une a estos puntos y sea $x_0 \in \Omega$. Entonces se tiene*

$$\|f(b)-f(a)-Df(x_0)(b-a)\| \le \|b-a\| \sup_{x\in S}\{\|Df(x)-Df(x_0)\|_{pq}\}.$$

DEMOSTRACION. Defina $g:\Omega \to \boldsymbol{R}^q$ para $x \in \Omega$ como

$$g(x) = f(x) - Df(x_0)(x).$$

Dado que $Df(x_0)$ es lineal, se sigue que $Dg(x) = Df(x) - Df(x_0)$ para $x \in \Omega$. Si se aplica el teorema del valor medio 40.5, se infiere que existe un punto $c \in S$ tal que

$$\begin{aligned}\|f(b)-f(a)-Df(x_0)(b-a)\| &= \|g(b)-g(a)\| \\ &\le \|Dg(c)(b-a)\| = \|(Df(c)-Df(x_0))(b-a)\| \\ &\le \|b-a\| \sup_{x\in S}\{\|Df(x)-Df(x_0)\|_{pq}\}.\end{aligned}$$

Q.E.D.

El siguiente resultado es el lema clave para los teoremas sobre transformaciones.

41.4 LEMA DE APROXIMACION. *Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y suponga que $f:\Omega \to \boldsymbol{R}^q$ pertenece a la clase $C^1(\Omega)$. Si $x_0 \in \Omega$ y $\varepsilon > 0$, entonces existe $\delta(\varepsilon) > 0$ tal que si $\|x_k - x_0\| \le \delta(\varepsilon)$, $k = 1, 2$, entonces $x_k \in \Omega$*

(41.2) $$\|f(x_1)-f(x_2)-Df(x_0)(x_1-x_2)\| \le \varepsilon\, \|x_1 - x_2\|.$$

DEMOSTRACION. Dado que $x \mapsto Df(x)$ es continua de Ω a $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$, entonces dada $\varepsilon > 0$ existe $\delta(\varepsilon) > 0$ tal que si $\|x - x_0\| < \delta(\varepsilon)$ entonces $x \in \Omega$ y $\|Df(x) - Df(x_0)\|_{pq} \le \varepsilon$. Ahora, suponga que x_1, x_2 satisfacen $\|x_k - x_0\| \le \delta(\varepsilon)$, por lo que el segmento de línea que une a x_1 y x_2 queda dentro de la bola cerrada con centro x_0 y radio $\delta(\varepsilon)$, y, por lo tanto, dentro de Ω. Se aplica ahora el lema 41.3 para obtener la conclusión deseada. Q.E.D.

El teorema de la transformación inyectiva

En seguida se probará que si f pertenece a la clase $C^1(\Omega)$ y si $Df(c)$ es inyectiva, entonces la restricción de f para una vecindad apropiada de c es una inyección.

El lector que esté familiarizado con el concepto de "rango" de una transformación lineal recordará que $L:\boldsymbol{R}^p \to \boldsymbol{R}^q$ es inyectiva si y sólo si el rango $(L) = p \le q$.

41.5 TEOREMA DE LA TRANSFORMACION INYECTIVA. *Suponga que $\Omega \subseteq \boldsymbol{R}^p$ es abierto, que $f:\Omega \to \boldsymbol{R}^q$ pertenece a la clase $C^1(\Omega)$,*

y que $L = Df(c)$ *es una inyección. Entonces, existe un número* $\delta > 0$ *tal que la restricción de f a* $B_\delta = \{x \in \boldsymbol{R}^p : \|x - c\| \leq \delta\}$ *es una inyección. Más aún, el inverso de la restricción* $f \mid B_\delta$ *es una función continua en* $f(B_\delta) \subseteq \boldsymbol{R}^q$ *a* $B_\delta \subseteq \boldsymbol{R}^p$.

DEMOSTRACION. Dado que la función lineal $L = Df(c)$ es una inyección, del corolario 22.7 se infiere que existe una $r > 0$ tal que

$$(41.3) \qquad r\,\|u\| \leq \|Df(c)(u)\| \quad \text{para } u \in \boldsymbol{R}^p.$$

Se aplica ahora el lema de aproximación 41.4 con $\varepsilon = \frac{1}{2}r$ para obtener un número $\delta > 0$ tal que si $\|x_k - c\| \leq \delta$, $k = 1, 2$, entonces

$$\|f(x_1) - f(x_2) - L(x_1 - x_2)\| \leq \tfrac{1}{2}r\,\|x_1 - x_2\|.$$

Si se aplica la desigualdad del triángulo al lado izquierdo de esta desigualdad se obtiene

$$\|L(x_1 - x_2)\| - \|f(x_1) - f(x_2)\| \leq \tfrac{1}{2}r\,\|x_1 - x_2\|.$$

Combinando esto con (41.3), en donde $u = x_1 - x_2$, se obtiene

$$(41.4) \qquad \tfrac{1}{2}r\,\|x_1 - x_2\| \leq \|f(x_1) - f(x_2)\|$$

para $x_k \in B_\delta$. Esto prueba que la restricción de f B_δ es una inyección; por lo tanto, esta restricción tiene una función inversa a la que se designará por medio de g. Si $y_k \in f(B_\delta)$, entonces existen puntos únicos $x_k = g(y_k)$ en B_δ tal que $y_k = f(x_k)$. Se infiere de (41.4) que

$$\|g(y_1) - g(y_2)\| \leq (2/r)\,\|y_1 - y_2\|,$$

por lo que $g = (f \mid B_\delta)^{-1}$ es uniformemente continua de $f(B_\delta)$ a $\boldsymbol{R}^p$.

Observe que g no está definida necesariamente en una vecindad de $f(c)$, es decir $f(c)$ no necesariamente es un punto interior de $f(B_\delta)$. Por este motivo, no se puede asegurar nada acerca de la diferenciabilidad de g. En seguida se demostrará un teorema de inversión más efectivo con hipótesis adicionales.

El teorema de transformación suprayectiva

El siguiente resultado va unido al teorema de transformación inyectiva. Este teorema, debido a L.M.Graves†, asegura que si f está en la clase $C^1(\Omega)$ y si para alguna $c \in \Omega$, la aplicación lineal $Df(c)$ es una suprayección de $\boldsymbol{R}^p$ sobre $\boldsymbol{R}^q$, entonces f aplica una vecindad apropiada de c sobre sobre una vecindad de $f(c)$. De modo que todo punto de $\boldsymbol{R}^q$ suficientemente cerca a $f(c)$ es la imagen bajo f de un punto cerca de c.

LAWRENCE M. GRAVES (1896-1973) nació en Kansas pero durante muchos años, como estudiante y profesor. Se le conoce principalmente por sus aportaciones al análisis funcional y al cálculo de variaciones.

El lector que esté familiarizado con el concepto del "rango" de una transformación lineal recordará que $L: \mathbf{R}^p \to \mathbf{R}^q$ es suprayectiva si y sólo si el rango $(L) = q \leq p$.

41.6 TEOREMA DE TRANSFORMACION SUPRAYECTIVA. *Sea $\Omega \subseteq \mathbf{R}^p$ abierto y suponga que $f: \Omega \to \mathbf{R}^q$ pertenece a la clase $C^1(\Omega)$. Suponga que para alguna $c \in \Omega$, la función lineal $L = Df(c)$ es una suprayección de $\mathbf{R}^p$ sobre $\mathbf{R}^q$. Entonces, existen números $m > 0$ y $\alpha > 0$ tales que si $y \in \mathbf{R}^q$ y $\|y - f(c)\| \leq \alpha/2m$, entonces existe una $x \in \Omega$ tal que $\|x - c\| \leq \alpha$ y $f(x) = y$.*

DEMOSTRACION. Dado que L es una suprayección, cada uno de los vectores de la base usual

$$e_1 = (1, 0, \dots, 0), \qquad e_2 = (0, 1, \dots, 0), \dots, e_q = (0, 0, \dots, 1)$$

en $\mathbf{R}^q$ es la imagen bajo L de algún vector en $\mathbf{R}^p$, digamos $u_1, u_2, \dots, u_q$. Ahora, sea $M: \mathbf{R}^q \to \mathbf{R}^p$ la función lineal que aplica e_j en u_j para $j = 1, 2, \dots, q$; es decir

$$M\left(\sum_{i=1}^{q} a_i e_i\right) = \sum_{i=1}^{q} a_i u_i.$$

Se sigue que $L \circ M$ es la aplicación identidad en $\mathbf{R}^q$; es decir $L \circ M(y) = y$ para toda $y \in \mathbf{R}^q$. Si se toma

$$m = \left\{\sum_{i=1}^{q} \|u_i\|^2\right\}^{1/2},$$

entonces aplicación de las desigualdades del triángulo y de Schwarz implican que si $y = \sum_{i=1}^{q} a_i e_i$, entonces

$$\begin{aligned} \|M(y)\| &\leq \sum_{i=1}^{q} |a_i| \, \|u_i\| \\ &\leq \left\{\sum_{i=1}^{q} |a_i|^2\right\}^{1/2} \left\{\sum_{i=1}^{q} \|u_i\|^2\right\}^{1/2} \\ &= m \, \|y\|. \end{aligned}$$

Por el lema de aproximación 41.4, existe un número $\alpha > 0$ tal que si $\|x_k - c\| \leq \alpha$, $k = 1, 2$, entonces $x_k \in \Omega$ y

$$\|f(x_1) - f(x_2) - L(x_1 - x_2)\| \leq \frac{1}{2m} \|x_1 - x_2\|. \tag{41.5}$$

Ahora, sea $B_\alpha = \{x \in \mathbf{R}^p : \|x - c\| \leq \alpha\}$ y suponga que $y \in \mathbf{R}^q$ es tal que $\|y - f(c)\| \leq \alpha/2m$. Se probará que existe un vector x con $x \in B_\alpha$ tal que $y = f(x)$.

Sea $x_0 = c$ y sea $x_1 = x_0 + M(y - f(c))$ de tal manera que $\|x_1 - x_0\| \leq m \, \|y - f(c)\| \leq \frac{1}{2}\alpha$, de donde

$$\|x_1 - x_0\| \leq \frac{\alpha}{2} \qquad \text{y} \qquad \|x_1 - c\| \leq \left(1 - \frac{1}{2}\right)\alpha.$$

Suponga que $c = x_0, x_1, \ldots, x_n$ se han elegido por inducción en $\boldsymbol{R}^p$ tales que

(41.6) $$\|x_k - x_{k-1}\| \leq \alpha/2^k, \qquad \|x_k - c\| \leq (1 - 1/2^k)\alpha,$$

para $k = 1, \ldots, n$. Se define ahora $x_{n+1}\ (n \geq 1)$ como

(41.7) $$x_{n+1} = x_n - M[f(x_n) - f(x_{n-1}) - L(x_n - x_{n-1})].$$

De (41.5) se infiere que

$$\|x_{n+1} - x_n\| \leq m\ \|f(x_n) - f(x_{n-1}) - L(x_n - x_{n-1})\|$$
$$\leq \frac{1}{2}\|x_n - x_{n-1}\|,$$

por lo que $\|x_{n+1} - x_n\| \leq \frac{1}{2}(\alpha/2^n) = \alpha/2^{n+1}$ y

$$\|x_{n+1} - c\| \leq \|x_{n+1} - x_n\| + \|x_n - c\|$$
$$\leq (\alpha/2^{n+1}) + (1 - 1/2^n)\alpha$$
$$= (1 - 1/2^{n+1})\alpha.$$

De donde (41.6) también queda probado para $k = n + 1$. Por lo tanto, se puede construir una sucesión (x_n) en B_α de esta manera. Si $m \geq n$, se tiene

$$\|x_n - x_m\| \leq \|x_n - x_{n+1}\| + \|x_{n+1} - x_{n+2}\| + \cdots + \|x_{m-1} - x_m\|$$
$$\leq \frac{\alpha}{2^{n+1}} + \frac{\alpha}{2^{n+2}} + \cdots + \frac{\alpha}{2^m} \leq \frac{\alpha}{2^n}.$$

Se deduce que (x_n) es una sucesión de Cauchy en $\boldsymbol{R}^p$ y por lo tanto converge a algún elemento x. Dado que $\|x_n - c\| \leq (1 - 1/2^n)\alpha$, se sigue que $\|x - c\| \leq \alpha$ de manera que $x \in B_\alpha$.

Como $x_1 - x_0 = M(y - f(c))$, se deduce que

$$L(x_1 - x_0) = L \circ M(y - f(c)) = y - f(x_0).$$

Más aún, por (41.7) se tiene

$$L(x_{n+1} - x_n) = -L \circ M[f(x_n) - f(x_{n-1}) - L(x_n - x_{n-1})]$$
$$= -\{f(x_n) - f(x_{n-1}) - L(x_n - x_{n-1})\}$$
$$= L(x_n - x_{n-1}) - [f(x_n) - f(x_{n-1})]$$

Por inducción se obtiene

$$L(x_{n+1} - x_n) = y - f(x_n),$$

por lo que se infiere que $y = \lim f(x_n) = f(x)$. Por lo tanto, todo punto y que

satisfaga $\|y - f(c)\| \leq \alpha/2m$ es la imagen bajo f de un punto $x \in \Omega$ con $\|x - c\| \leq \alpha$.

41.7 TEOREMA DE LA TRANSFORMACION ABIERTA. Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y suponga que $f:\Omega \to \boldsymbol{R}^q$ pertenece a la clase $C^1(\Omega)$. Si para cada $x \in \Omega$ la derivada $Df(x)$ es suprayección y si $G \leqslant \Omega$ es abierto, entonces $f(G)$ es abierto en $\boldsymbol{R}^q$.

DEMOSTRACION. Si $b \in f(G)$, entonces existe un punto $c \in G$ tal que $f(c) = b$. Del teorema de aplicación suprayectiva 41.6 aplicado a $f \mid G$ se infiere que existe $\beta > 0$ tal que si $\|y - b\| \leq \beta$ entonces existe una $x \in G$ tal que si $y = f(x)$. Por lo tanto, $f(g)$ es abierto en $\boldsymbol{R}^q$.

El teorema de inversión

Ahora se combinarán los dos teoremas de aplicación para el caso $p = q$. En este caso se va a suponer que a derivada $Df(c)$ es una biyección. Este es el caso si y sólo si la derivada $Df(c)$ tiene inverso, que a su vez es cierto si y sólo si el determinante jacobiano

$$J_f(c) = \det [D_j f_i(c)] = \det [f_{i,j}(c)]$$

es distinto de cero.

El lector familiarizado con el concepto del "rango" de una transformación lineal recordará que $L:\boldsymbol{R}^p \to \boldsymbol{R}^q$ es biyectiva si y sólo si el rango $(L) = p = q$.

De la continuidad de las derivadas parciales y del determinante se infiere que si $Df(a)$ es invertible, entonces $Df(x)$ es invertible para x suficientemente cerca de c.

41.8 TEOREMA DE INVERSION. *Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y suponga que $f:\Omega \to \boldsymbol{R}^p$ pertenece a la clase $C^1(\Omega)$. Si $c \in \Omega$ es tal que $Df(c)$ es una biyección, entonces existe una vecindad abierta U de c tal que $V = f(U)$ es una vecindad abierta de $f(c)$ y la restricción de f a U es una biyección sobre V con inverso continuo. Más aún, g pertenece a la clase $C^1(V)$* y

$$Dg(y) = [Df(g(y))]^{-1} \qquad \text{para } y \in V.$$

DEMOSTRACION. Por hipótesis $L = Df(c)$ es inyectiva; por lo tanto, el corolario 22.7 implica que existe $r > 0$ tal que

$$2r\|z\| \leq \|Df(c)(z)\| \qquad \text{para } z \in \boldsymbol{R}^p.$$

Dado que f está en la clase $C^1(\Omega)$, hay una vecindad de c en la que $Df(x)$ es invertible y satisface

(41.8) $$r\,\|z\| \le \|Df(x)(z)\| \qquad \text{para } z \in \boldsymbol{R}^p.$$

De nuevo la atención se limita a una vecindad U de c en la que f es inyectiva y que está contenida en un bola con centro c y radio α (como en el teorema de transformación suprayectiva 41.6). Entonces, $V = f(U)$ es una vecindad de $f(c)$ y de los teoremas de transformación anteriores se infiere que la restricción $f\,|\,U$ tiene una función continua inversa $g: V \to \boldsymbol{R}^p$.

Queda por demostrar que g es diferenciable en un punto arbitrario $y_1 \in V$. Sea $x_1 = g(y_1) \in U$; dado que es diferenciable en x_1, se deduce que si $x \in U$, entonces

$$f(x) - f(x_1) - Df(x_1)(x - x_1) = \|x - x_1\|\, u(x),$$

en donde $\|u(x)\| \to 0$ conforme $x \to x_1$. Si M_1 s el inverso de la función lineal $Df(x_1)$, entonces

$$\begin{aligned} x - x_1 &= M_1[Df(x_1)(x - x_1)] \\ &= M_1[f(x) - f(x_1) - \|x - x_1\|\, u(x)]. \end{aligned}$$

Si $x \in U$, entonces $x = g(y)$ para alguna $y = f(x) \in V$; además $y_1 = f(x_1)$, y esta ecuación se puede escribir en la forma

$$g(y) - g(y_1) - M_1(y - y_1) = -\|x - x_1\|\, M_1(u(x)).$$

Dado que $Df(x_1)$ es inyectiva, como en la demostración del teorema de transformación inyectiva 41.5 se infiere que

$$\|y - y_1\| = \|f(x) - f(x_1)\| \ge \tfrac{1}{2} r\, \|x - x_1\|$$

siempre que y esté lo suficientemente cerca de y_1. Además, de (41.8) se deduce que $\|M_1(u)\| \le (1/r)\,\|u\|$ para toda $u \in \boldsymbol{R}^p$. Por lo tanto, se tiene

$$\|g(y) - g(y_1) - M_1(y - y_1)\| \le (2/r^2)\, \|u(x)\|\, \|y - y_1\|.$$

Ahora, como $y \to y_1$, entonces $x = g(y) \to g(y_1) = x_1$ y $\|u(x)\| \to 0$. Se concluye, por lo tanto, que $Dg(y_1)$ existe es igual a $M_1 = (Df(x_1))^{-1}$.

El hecho de que g pertenezca a la clase $C^1(V)$ se sigue de la relación $Dg(y) = [Df(g(y))]^{-1}$ para $y \in V$, y de la continuidad de las aplicaciones

$$y \mapsto g(y), \qquad x \mapsto Df(x), \qquad L \mapsto L^{-1}$$

de $V \to U$, $U \to \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^p)$, y $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^p) \to \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^p)$, respectivamente. (Véase el ejercicio 4 .L) Q.E.D.

Funciones implicitas

Suponga que F es una función definida en un subconjunto de $\boldsymbol{R}^p \times \boldsymbol{R}^q$ a (Si se hace la identificación obvia de $\boldsymbol{R}^p \times \boldsymbol{R}^q$ con $\boldsymbol{R}^{p+q}$, entonces no es

necesario definir lo que significa que F sea continua o que sea diferenciable en un punto o que esté en la clase C^1 en un conjunto.) Suponga que F aplica al punto (a, b) en el vector cero de5 $\boldsymbol{R}^q$.El problema de las funciones implícitas consiste en resolver la ecuación

$$F(x, y) = 0$$

para un elemento (digamos, y) en términos del otro en el sentido de que se encuentra una función φ definida en un subconjunto de $\boldsymbol{R}^p$ con valores en $\boldsymbol{R}^q$ tal que $b = \varphi(a)$ y

$$F(x, \varphi(x)) = 0$$

para toda x en el dominio de φ. Se supone que F es continua en una vecindad de (a, b), se desea a la conclusión de que la "función solución" φ es continua en una vecindad de a. Tal vez no sea ninguna sorpresa para el lector el que se deba suponer que F pertenece a la clase C^1 en una vecindad de (a, b); sin embargo, aun esta hipótesis no es suficiente para garantizar la existencia y unicidad de una función solución continua φ definida en una vecindad de a.

De hecho, si $p = q = 1$, la función dada por $F(x, y) = x^2 - y^2$ tiene dos funciones solución continuas $\varphi_1(x) = x$ y $\varphi_2(x) = -x$ correspondientes al punto (0,0). También tiene soluciones discontinuas, tales como

$$\begin{aligned} \varphi_3(c) &= x, && x \text{ racional}, \\ &= -x, && x \text{ irracional}. \end{aligned}$$

La función $G(x, y) = x - y^2$ tiene dos funciones solución continuas correspondientes a (0,0), pero ninguna de ellas está definida en una vecindad del punto $x = 0$. Para dar un ejemplo más complejo, la función $H: \boldsymbol{R}^2 \to \boldsymbol{R}$ definida por

$$\begin{aligned} H(x, y) &= x, && y = 0, \\ &= x - y^3 \operatorname{sen}\left(\frac{1}{y}\right), && y \neq 0, \end{aligned}$$

pertenece a la clase C^1 en una vecindad (0,0), pero no hay ninguna función solución continua definida en una vecindad de $x = 0$.

En estos tres ejemplos la derivada parcial con respecto a y desaparece en el punto considerado. En el caso $p = q = 1$, la afirmación adicional necesaria para asegurar la existencia y unicidad de la función solución es que esta derivada parcial sea distinta de cero. En el caso general, se observa que $DF(a, b)$ es una función lineal continua de $\boldsymbol{R}^p \times \boldsymbol{R}^q$ a $\boldsymbol{R}^q$ e induce una función lineal continua $L_2: \boldsymbol{R}^q \to \boldsymbol{R}^q$ definida por

$$L_2(v) = DF(a, b)(0, v)$$

para ver $v \in \boldsymbol{R}^q$. En un sentido muy razonable, L_2 es la "derivada parcial" de F con respecto a $y \in \boldsymbol{R}^q$ en el punto (a, b). El supuesto adicional que se va a imponer es que L_2 sea invertible.

Se desea interpretar este problema ahora en términos de coordenadas. Si $x = (x_1, \ldots, x_p)$ y $y = (y_1, \ldots, y_q)$, entonces la ecuación $F(x, y) = 0$ toma la forma de q ecuaciones en los $p+q$ argumentos $x_1, \ldots, x_p, y_1, \ldots, y_q$ dadas por

$$
\text{(41.9)} \qquad \begin{array}{c} f_1(x_1, \ldots, x_p, y_1, \ldots, y_q) = 0, \\ \cdots\cdots\cdots\cdots\cdots \\ f_q(x_1, \ldots, x_p, y_1, \ldots, y_q) = 0. \end{array}
$$

Por conveniencia, suponga que $a = 0$ y $b = 0$ de tal manera que este sistema se satisface para $x_1 = 0, \ldots, x_p = 0, y_1 = 0, \ldots, y_q = 0$, se desea resolver para las y_j en términos de las x_i al menos cuando las $|x_i|$ son suficientemente pequeñas. Si las funciones f_j son lineales, entonces la condición de solubilidad es que el determinante de coeficientes de las y_j deba ser distinto de cero. Si las funciones f_j no son lineales, entonces la condición es que el determinante jacobiano

$$
\frac{\partial(f_1, \ldots, f_q)}{\partial(y_1, \ldots, y_q)}(a, b) \neq 0.
$$

Cuando este es el caso, hay funciones φ_j, $j = 1, \ldots, q$, definidas y continuas cerca de $a = 0$ tales que si se sustituye

$$
\begin{array}{c} y_1 = \varphi_1(x_1, \ldots, x_p), \\ \cdots\cdots\cdots \\ y_q = \varphi_q(x_1, \ldots, x_p), \end{array}
$$

en el sistema (41.9), se obtiene una identida en las x_i.

41.9 TEOREMA DE LA FUNCION IMPLICITA. *Sea* $\Omega \subseteq \mathbf{R}^p \times \mathbf{R}^q$ *abierto y sea* $(a, b) \in \Omega$. *Suponga que* $F: \Omega \to \mathbf{R}^q$ *pertenece a la clase* $C^1(\Omega)$, *que* $F(a, b) = 0$, *y que la aplicación lineal definida por*

$$
L_2(v) = DF(a, b)(0, v), \qquad v \in \mathbf{R}^q,
$$

es una biyección de $\mathbf{R}^q$ *sobre* $\mathbf{R}^q$.

(a) Entonces existe una vecindadad abierta W *de* $a \in \mathbf{R}^p$ *y una función única* $\varphi: W \to \mathbf{R}^q$ *perteneciente a la clase* $C^1(W)$ *tal que* $b = \varphi(a)$ *y*

$$
F(x, \varphi(x)) = 0 \text{ para toda } x \in W.
$$

(b) Existe una vecindad abierta U *de* (a, b) *en* $\mathbf{R}^p \times \mathbf{R}^q$ *tal que el par* $(x, y) \in U$ *satisface* $F(x, y) = 0$ *si y sólo si* $y = \varphi(x)$ *para* $x \in W$.

DEMOSTRACION. Sin perder generalidad se puede suponer que $a = 0$ y $b = 0$. Suponga que $H: \Omega \to \mathbf{R}^p \times \mathbf{R}^q$ está definida como

$$
H(x, y) = (x, F(x, y)) \qquad \text{para } (x, y) \in \Omega.
$$

Se sigue fácilmente (véase el ejercicio 39.V) que H pertenece a la clase $C^1(\Omega)$ y que

$$DH(x, y)(u, v) = (u, DF(x, y)(u, v))$$

para $(x, y) \in \Omega$ y $(u, v) \in \boldsymbol{R}^p \times \boldsymbol{R}^q$. Se asegura ahora que $DH(0,0)$ es invertible en $\boldsymbol{R}^q \times \boldsymbol{R}^q$. De hecho, si se toma $L_1 \in \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ definida como

$$L_1(u) = DF(0, 0)(u, 0) \qquad \text{para } u \in \boldsymbol{R}^p;$$

entonces el hecho de que $DF(0, 0)(u, v) = L_1(u) + L_2(v)$ prueba que el inverso de $DH(0,0)$ es la aplicación lineal K en $\boldsymbol{R}^p \times \boldsymbol{R}^q$ definida por

$$K(x, z) = (x, L_2^{-1}[z - L_1(x)]).$$

Por lo tanto, del teorema de inversión 41.8 se infiere que hay una vecindad abierta U de $(0, 0) \in \boldsymbol{R}^p \times \boldsymbol{R}^q$ tal que $V = H(U)$ es una vecindad abierta de $(0, 0) \in \boldsymbol{R}^p \times \boldsymbol{R}^q$ y la restricción de H a U es una biyección sobre V con un inverso continuo $\Phi : V \to U$ que pertenece a la clase $C^1(V)$ y con $\Phi(0, 0) = (0, 0)$. Ahora Φ es la forma

$$\Phi(x, z) = (\varphi_1(x, z), \varphi_2(x, z)) \qquad \text{para } (x, z) \in V$$

en donde $\varphi_1 : V \to \boldsymbol{R}^p$ y $\varphi_2 : V \to \boldsymbol{R}^q$. Dado que

$$\begin{aligned}(x, z) = H \circ \Phi(x, z) &= H[\varphi_1(x, z), \varphi_2(x, z)] \\ &= [\varphi_1(x, z), F(\varphi_1(x, z), \varphi_2(x, z))],\end{aligned}$$

se deduce que $\varphi_1(x, z) = x$ para toda $(x, z) \in V$. por lo que Φ toma la forma más simple

$$\Phi(x, z) = (x, \varphi_2(x, z)) \qquad \text{para } (x, z) \in V.$$

Ahora, si $P : \boldsymbol{R}^p \times \boldsymbol{R}^q \to \boldsymbol{R}^q$ está definida como $P(x, z) = z$, entonces P es lineal y continua y $\varphi_2 = P \circ \Phi$; por lo tanto, φ_2 pertenece a la clase $C^1(V)$ y se tiene

$$z = F(x, \varphi_2(x, z)) \qquad \text{para } (x, z) \in V.$$

Sea ahora $W = \{x \in \boldsymbol{R}^p : (x, 0) \in V\}$ de manera que W es una vecindad abierta de 0 en $\boldsymbol{R}^p$, y defina $\varphi(x) = \varphi_2(x, 0)$ para $x \in W$. es evidente que $\varphi(0) = 0$, y de la fórmula anterior se infiere que

$$F(x, \varphi(x)) = 0 \qquad \text{para } x \in W.$$

Más aún, $D\varphi(x)(u) = D\varphi_2(x, 0)(u, 0)$ para $x \in W$, $u \in \boldsymbol{R}^p$, por lo que se concluye que φ pertenece a la clase $C^1(W)$. Esto prueba la parte *(a)*.

Para completar la demostración de la parte *(b)*, suponga que $(x, y) \in U$ satisface $F(x, y) = 0$. Entonces $H(x, y) = (x, F(x, y)) = (x, 0) \in V$ por lo que se sigue $x \in W$. Además, $(x, y) = \Phi(x, 0) = (x, \varphi_2(x, 0)) = (x, \varphi(x))$ de tal manera que $y = \varphi(x)$. Q.E.D.

Algunas veces es útil tener una fórmula explícita para la derivada de φ. Para poder dar esto, es conveniente introducir el concepto de las **derivadas**

parciales de bloque de F. Si $(x, y) \in \Omega$, la derivada parcial de bloque $D_{(1)}F(x, y)$ es la función lineal que aplica $\mathbf{R}^q \to \mathbf{R}^q$ dada por

$$D_{(1)}F(x, y)(u) = Df(x, y)(u, 0) \quad \text{para } u \in \mathbf{R}^p,$$

y la derivada parcial de bloque $D_{(2)}F(x, y)$ es la función lineal que aplica $\mathbf{R}^p \to \mathbf{R}^q$ dada por

$$D_{(2)}F(x, y)(v) = DF(x, y)(0, v) \quad \text{para } v \in \mathbf{R}^q.$$

Dado que $(u, v) = (u, 0) + (0, v)$ es claro que

$$(41.10) \qquad DF(x, y)(u, v) = D_{(1)}F(x, y)(u) + D_{(2)}F(x, y)(v).$$

Observe que las transformaciones L_1 y L_2 de la demostración anterior son $D_{(1)}F(0, 0)$ y $D_{(2)}F(0, 0)$, respectivmente.

41.10 COROLARIO. *Con la hipótesis del torema, existe una* $\gamma > 0$ *tal que si* $\|x - a\| < \gamma$, *entonces la derivada de* φ *en* x *es el elemento de* $\mathcal{L}(\mathbf{R}^p, \mathbf{R}^q)$ *dado por*

$$(41.11) \qquad D\varphi(x) = -[D_{(2)}F(x, \varphi(x))]^{-1} \circ [D_{(1)}F(x, \varphi(x))].$$

DEMOSTRACION. Suponga que $K : W \to \mathbf{R}^p \times \mathbf{R}^q$ está definida por

$$K(x) = (x, \varphi(x)) \quad \text{para } x \in W.$$

Entonces, dado que $F \circ K(x) = F(x, \varphi(x)) = 0$, se tiene $F \circ K : W \to \mathbf{R}^q$ es una función constante. Además, como fácilmente se puede ver que

$$DK(x)(u) = (u, D\varphi(x)(u)) \quad \text{para } u \in \mathbf{R}^p,$$

de la regla de la cadena 40.2 aplicada a la función constante $F \circ K$ se infiere que

$$0 = D(F \circ K)(x) = DF(K(x)) \circ DK(x).$$

Usando (41.10) se tiene

$$DF(x, \varphi(x))(u, v) = D_{(1)}F(x, \varphi(x))(u) + D_{(2)}F(x, \varphi(x))(v).$$

De aquí se deduce que si $u \in \mathbf{R}^p$, entonces

$$\begin{aligned} 0 = DF(x, \varphi(x))(u) &= D_{(1)}F(x, \varphi(x))(u) + D_{(2)}F(x, \varphi(x))(D\varphi(x)(u)) \\ &= D_{(1)}F(x, \varphi(x))(u) + [D_{(2)}F(x, \varphi(x)) \circ D\varphi(x)](u). \end{aligned}$$

Por lo que se tiene

$$0 = D_{(1)}F(x, \varphi(x)) + D_{(2)}F(x, \varphi(x)) \circ D\varphi(x)$$

para toda $x \in W$. Por hipótesis, $L_2 = D_{(2)}F(a, b)$ es invertible. Dado que φ y F son continuas, existe una $\gamma > 0$ tal que si $\|x - a\| < \gamma$, entonces $D_{(2)}F(x, \varphi(x))$

también es invertible. Por lo tanto, la ecuación (41.11) se deduce de la ecuación anterior. Q.E.D.

Puede ser útil interpretar la fórmula (41.11) en términos de matrices. Suponga que se tiene un sistema de q ecuaciones en $p+q$ argumentos dado por (41.9). Como ya se ha dicho antes, la hipótesis del teorema de la función implícita requiere que la matriz

$$\begin{bmatrix} f_{1,p+1} & \cdots & f_{1,p+q} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ f_{q,p+1} & \cdots & f_{q,p+q} \end{bmatrix}$$

sea invertible en el punto (a, b). (Recuerde que $f_{i,j}$ denota a la derivada parcial de f_i con respecto al *j-ésimo* argumento.) En este caso la derivada de la función solución φ en un punto x está dada por

$$-\begin{bmatrix} f_{1,p+1} & \cdots & f_{1,p+q} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ f_{q,p+1} & \cdots & f_{q,p+q} \end{bmatrix}^{-1} \begin{bmatrix} f_{1,1} & \cdots & f_{1,p} \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ f_{q,1} & \cdots & f_{q,p} \end{bmatrix},$$

en donde se entiende que ambas matrices están calculadas en el punto $(x, \varphi(x))$ cercano a (a, b).

Los teoremas de parametrización y de rango

El teorema de la función implícita se puede considerar como aquel que da las condiciones bajo las cuales la "curva de nivel"

$$C = \{(x, y) \in \boldsymbol{R}^p \times \boldsymbol{R}^q : F(x, y) = 0\}$$

que pasa por el punto (a, b) se puede parametrizar, cuando menos localmente, como la *gráfica* en $\boldsymbol{R}^p \times \boldsymbol{R}^q$ de alguna función definida en una vecindad W de $a \in \boldsymbol{R}^p$ a $\boldsymbol{R}^q$; es decir

$$C = \{(x, \varphi(x)) : x \in W\}.$$

Se estudiará ahora otro teorema que ofrece condiciones bajo las cuales la imagen de una función que transforma un subconjunto abierto de $\boldsymbol{R}^p$ hacia $\boldsymbol{R}^q$ se puede parametrizar por medio de una función φ definida en un conjunto abierto en un espacio de menor dimensión.

Al estudiar este teorema será necesario usar algunos hechos elementales pero importantes del álgebra lineal quepueden ser ya familiares para el lector.† Recuerde que si $L : \boldsymbol{R}^p \to \boldsymbol{R}^q$ es una transformación lineal entonces la **imagen** R_L de L es el subespacio de $\boldsymbol{R}^q$ dado por

† Para más detalles consultar los libros de Hoffman y Kunze o Finkbeiner que aparecen en la lista de referencias

$$R_L = \{L(x) : x \in \boldsymbol{R}^p\},$$

y el **espacio nulo** (o el **kernel**). N_L de L es el subespacio de $\boldsymbol{R}^p$ dado por

$$N_L = \{x \in \boldsymbol{R}^p : L(x) = 0\}.$$

La dimensión $r(L)$ de R_L se llama el **rango** de L y la dimensión $n(L)$ de N_L se llama la **nulidad** de L. (De modo que el rango de L es el número de vectores linealmente independientes en $\boldsymbol{R}^q$ necesarios para generar la imagen R_L, y la nulidad de L es el número de vectores linealmente independientes en $\boldsymbol{R}^p$ necesarios para generar el espacio nulo N_L.) Queda como ejercicio demostrar que si $\{u_1, \ldots, u_n\}$ (en donde $n = n(L)$) es un conjunto linealmente independiente de vectores en $\boldsymbol{R}^p$ que generan N_L al que se le agregan p-n vectores $u_{n+1}, \ldots, u_p$ para obtener una base para $\boldsymbol{R}^p$, entonces el conjunto $\{L(u_{n+1}), \ldots, L(u_p)\}$ es un conjunto linealmente independiente de vectores en $\boldsymbol{R}^q$ que generan R_L. Por lo tanto, se sigue que $p = n(L) + r(L)$; de donde: *la dimensión del dominio de L es igual a la suma de la nulidad y el rango de L.*

Si L se representa por una matriz $q \times p$ como en (23.1), entonces se puede probar que el rango de L es el máximo número r tal que exista al menos una submatriz $r \times r$ con determinante distinto de cero. El teorema de parametrización asegura que si f es una aplicación C^1 de un conjunto abierto $\Omega \subseteq \boldsymbol{R}^p$ hacia $\boldsymbol{R}^q$ tal que $Df(x)$ tenga rango igual a r para toda $x \in \Omega$ y si $f(a) = b \in \boldsymbol{R}^q$ para alguna $a \in \Omega$, entonces hay una vecindad V de a tal que la restricción de f a V se pueda dar como una aplicación C^1 φ definida en una vecindad en $\boldsymbol{R}^r$.

41.11 TEOREMA DE PARAMETRIZACION. *Sea* $\Omega \subseteq \boldsymbol{R}^p$ *abierto y suponga que* $f : \Omega \to \boldsymbol{R}^p$ *pertenece a la clase* $C^1(\Omega)$. *Suponga que* $Df(x)$ *tiene rango* r *para toda* $x \in \Omega$ *y sea* $f(a) = b \in \boldsymbol{R}^q$ *para alguna* $a \in \Omega$. *Entonces*

(i) existe una vecindad abierta $V \subseteq \Omega$ *de a y una función* $\alpha : V \to \boldsymbol{R}^r$ *en la clase* $C^1(V)$, *y*

(ii) existe un conjunto abierto $W \subseteq \boldsymbol{R}^r$ *y funciones* $\beta : W \to \boldsymbol{R}^p$ *y* $\varphi : W \to \boldsymbol{R}^q$, *tales que*

(iii) $f(x) = \varphi \circ \alpha(x)$ *para toda* $x \in V$, *y* $\varphi(t) = f \circ \beta(t)$ *para toda* $t \in W$.

DEMOSTRACION. Sin perder generalidad se puede suponer que $a = 0 \in \boldsymbol{R}^p$ y $b = 0 \in \boldsymbol{R}^q$.

Sea $L = Df(0)$ de manera que $L : \boldsymbol{R}^p \to \boldsymbol{R}^q$ tega rango r y sea $\{x_1, \ldots, x_p\}$ una base en $\boldsymbol{R}^p$ tal que $\{x_{r+1}, \ldots, x_p\}$ genere al espacio nulo de L. Se toma X_1 como el subespacio generador por $\{x_1, \ldots, x_r\}$ y $X_2 = N_L$ como el subespacio generado por $\{x_{r+1}, \ldots, x_p\}$. Como se mencionó poco antes, se infiere que $Y_1 = R_L$ está generado por $\{y_1 = L(x_1), \ldots, y_r = L(x_r)\}$. Se elige $\{y_{r+1}, \ldots, y_q\}$ tal que $\{y_1, \ldots, y_q\}$ sea una base para $\boldsymbol{R}^q$ y suponga que Y_2 es el subespacio generado por $\{y_{r+1}, \ldots, y_q\}$.

Se sigue que todo vector $x \in \boldsymbol{R}^p$ tiene una representación única de la forma $x = c_1x_1 + \cdots + c_px_p$. Sean P_1 y P_2 las transformaciones lineales en $\boldsymbol{R}^p$ definidas como

$$P_1(x) = \sum_{j=1}^{r} c_j x_j, \qquad P_2(x) = \sum_{j=r+1}^{p} c_j x_j.$$

Es claro que la imagen de P_j es igual a X_j, $j = 1, 2$. De manera análoga se toman Q_1 y Q_2 como transformaciones lineales en $\boldsymbol{R}^q$ definidas para $y = c_1 y_1 + \cdots + c_q y_q$ como

$$Q_1(y) = \sum_{j=1}^{r} c_j y_j, \qquad Q_2(y) = \sum_{j=r+1}^{q} c_j y_j.$$

Claramente, la imagen de Q_j es Y_j, $j = 1, 2$.

Si L_1 es la restricción de L a X_1, entonces L_1 es una biyección de X_1 sobre Y_1; sea $A : Y_1 \to X_1$ el inverso de L_1. se puede ver que, $A \circ L(x) = x$ para toda $x \in X_1$ y $L \circ A(y) = y$ para toda $y \in Y_1$. se define ahora u de $\Omega \subseteq \boldsymbol{R}^p$ a $\boldsymbol{R}^p$ como

(41.12) $$u(x) = A \circ Q_1 \circ f(x) + P_2(x), \qquad x \in \Omega,$$

de tal manera que $u(0) = 0$, u aplica $X_1 \cap \Omega$ en X_1, y

$$Du(x) = A \circ Q_1 \circ Df(x) + P_2, \qquad x \in \Omega;$$

de donde, u pertenece a la clase $C^1(\Omega)$. Puesto que fácilmente se puede ver que $Du(0)$ es la transformación identidad en $\boldsymbol{R}^p$, del teorema de inversión 41.8 se infiere que existe una vecindad abierta U de $a = 0$ tal que $U' = u(U)$ es una vecindad abierta de 0 y que la restricción de u a U es una biyección sobre U' con inverso $w = u^{-1} : U' \to \boldsymbol{R}^p$ que pertenece a la clase $C^1(U')$. Además, reemplazando U y U' por conjuntos más pequeños también se puede suponer que U' es convexo (es decir, contiene al segmento de línea que une a cualesquiera dos de sus puntos).

Ahora, suponga que $g : U' \to \boldsymbol{R}^q$ se define como

$$g(z) = f(w(z)), \qquad z \in U' \subseteq \boldsymbol{R}^p.$$

Está claro que g pertenece a la clase $C^1(U')$ y

$$Dg(z) = Df(w(z)) \circ Dw(z), \qquad z \in U'.$$

Dado que $Df(x)$ tiene rango r para toda $x \in \Omega$ y $Dw(z)$ es invertible para $x \in U'$, entonces se sigue de un teorema de álgebra lineal que $Dg(z)$ tiene rango r para toda $z \in U'$. Por otro lado,

$$\begin{aligned} g(z) &= (Q_1 + Q_2) \circ f(w(z)) \\ &= Q_1 \circ f(w(z)) + Q_2 \circ f(w(z)). \end{aligned}$$

Dado que $w = u^{-1}$, de (41.12) se infiere que

$$z = u(w(z)) = A \circ Q_1 \circ f(w(z)) + P_2(w(z)), \qquad z \in U'.$$

Pero como $L \circ A \circ Q_1 = Q_1$ en $\boldsymbol{R}^q$ y $L \circ P_2 = 0$ en $\boldsymbol{R}^p$, se tiene

(41.13) $$L(z) = Q_1 \circ f(w(z)) = Q_1 \circ g(z),$$

de donde $L = Q_1 \circ Dg(z)$ para $z \in U'$. por lo tanto, si $z \in U'$, el operador Q_1 aplica a la imagen de $Dg(z)$ (que tiene dimensión r) sobre la imagen de L (que también tiene dimensión r). Se sigue que Q_1 es inyectiva en la imagen de $Dg(z)$ para $z \in U'$; por lo tanto, si $z \in U'$ y $x \in \boldsymbol{R}^p$ es tal que $L(x) = 0$, entonces $Dg(z)(x) = 0$. en consecuencia, si $z \in U'$ y $z_2 \in X_2 = N_L$, se infiere que $Dg(z)(z_2) = 0$.

Se demostrará ahora que $g : U' \to \boldsymbol{R}^q$ depende sólo de $z_1 \in X_1$ en el sentido de que si $z \in U'$ y $z_2 \in X_2$ son tales que $z + z_2 \in U'$; entonces, $g(z + z_2) = g(z)$. Para ver esto se aplica el teorema del valor medio 40.5 para deducir que existe un punto z_0 en el segmento de línea que une a z y $z + z_2$ (y por lo tanto en U') tal que

$$0 \leq \|g(z + z_2) - g(z)\| \leq \|Dg(z_0)(z_2)\| = 0;$$

por lo tanto, $g(z + z_2) = g(z)$, como se quería.

Ahora ya se pueden definir las transformaciones α, β, φ. Sea $C : \boldsymbol{R}^r \to \boldsymbol{R}^p$ la transformación lineal que aplica a los elementos de la base usual $e_1, \ldots, e_r$ de $\boldsymbol{R}^r$ en los vectores $x_1, \ldots, x_r$ que forman una base para X_1. por lo tanto, C es una biyección de $\boldsymbol{R}^r$ sobre X_1 y entoces $C^{-1} : X_1 \to \boldsymbol{R}^r$ existe. Sea $W = C^{-1}(U') = C^{-1}(U' \cap X_1)$, de tal manera que $W \subseteq \boldsymbol{R}^r$ es una vecindad abierta de 0 en $\boldsymbol{R}^r$ y sea $V \subseteq U$ una vecindad abierta de $a = 0$ tal que $P_1 \circ u(V) \subseteq U'$. Se definen ahora $\alpha : V \to \boldsymbol{R}^r$ y $\beta : W \to \boldsymbol{R}^p$ como

(41.14) $$\alpha(x) = C^{-1} \circ P_1 \circ u(x), \qquad \beta(t) = w \circ C(t)$$

para $x \in V$ y $t \in W$. es claro que α pertenece a la clase $C^1(V)$ y $\alpha(V) \subseteq W$, y que β pertenece a la clase $C^1(W)$ y $\beta(W) \subseteq U$. Se define $\varphi : W \to \boldsymbol{R}^q$ para $t \in W$ como

$$\varphi(t) = g \circ C(t),$$

por lo que se sigue que

$$\varphi(t) = (f \circ w) \circ C(t) = f \circ \beta(t).$$

Más aún, si $x \in V$, entonces

$$f(x) = f(w \circ u(x)) = (f \circ w) \circ u(x) = g \circ u(x);$$

sin embargo, se ha visto que $g \circ u(x) = g \circ P_1 \circ u(x)$, de manera que

$$\begin{aligned} f(x) = g \circ u(x) &= g \circ (C \circ C^{-1}) \circ (P_1 \circ u)(x) \\ &= (g \circ C) \circ (C^{-1} \circ P_1 \circ u)(x) \\ &= \varphi \circ \alpha(x). \end{aligned}$$

Por lo tanto, $f(x) = \varphi \circ \alpha(x)$ para toda $x \in V$. Q.E.D.

Durante esta construcción, se ha establecido en realidad un poco más de información. En este corolario se hace uso de la notación que se desarrolló en la demostración del teorema.

41.12 COROLARIO. *(a) La aplicación* $\varphi : W \to \mathbf{R}^q$ *es de la forma* $\varphi_1 +$ *en donde* φ_2, φ_1 *es la restricción a* W *de la aplicación lineal de* $\mathbf{R}^r \to \mathbf{R}^q$ *que manda* $e_j \in \mathbf{R}^r$ *hacia* $y_j = L(x_j)$, $j = 1, \ldots, r$, *y en donde* $\varphi_2(W) \subseteq Y_2$.

(b) *Si* $t \in W$, *entonces* $\alpha \circ \beta(t) = t$.

(c) *Si* $x \in U \cap X_1$, *entonces* $x \in V$ *y* $\beta \circ \alpha(x) = x$.

DEMOSTRACION. *(a)* Dado que $g = Q_1 \circ g + Q_2 \circ g$, se infiere de (41.13) que $g = L + Q_2 \circ g$. Por lo tanto, a partir de la definición de φ, se tiene $\varphi = L \circ C + Q_2 \circ g \circ C$, que es de la forma establecida en *(a)*.

(b) Si $t \in W$, entonces $x = \beta(t) = w \circ C(t) \in U$ tiene la propiedad de que $u(x) = u \circ w \circ C(t) = C(t) \in U' \cap X_1$; por lo tanto $P_1 \circ u(x) = C(t) \in U'$ de modo que $x \in V$ y

$$\alpha(x) = C^{-1} \circ P_1 \circ u(x) = C^{-1} \circ C(t) = t,$$

que prueba la afirmación *(b)*.

(c) Si $x \in \Omega \cap X_1$, entonces, de (41.12) y del hecho de que $P_2(x) = 0$, se deduce que $u(x) \in X_1$. Por lo tanto, si $x \in U \cap X_1$ se sigue que $P_1 \circ u(x) = u(x) \in U' \cap X_1$ de tal manera que $x \in V$. Más aún,

$$\begin{aligned} \beta \circ \alpha(x) &= (w \circ C) \circ (C^{-1} \circ P_1 \circ u)(x) \\ &= w \circ C \circ C^{-1} \circ u(x) = w \circ u(x) = x. \end{aligned}$$

Q.E.D.

Se puede usar ahora el resultado del teorema de parametrización para demostrar el teorema del rango.

41.13 TEOREMA DEL RANGO. *Sea* $\Omega \subseteq \mathbf{R}^p$ *abierto y suponga que* $f : \Omega \to \mathbf{R}^q$ *pertenece a la clase* $C^1(\Omega)$. *Suponga que* $Df(x)$ *tiene rango* r *para toda* $x \in \Omega$ *y sea* $f(a) = b \in \mathbf{R}^q$ *para alguna* $a \in \Omega$. *Entonces*

(i) existen vecindades abiertas V *de* a *y* V' *de* 0 *en* $\mathbf{R}^p$ *y una función* $\sigma : V \to V'$ *en la clase* $C'(V)$ *que tiene un inverso* $\sigma^{-1} : V' \to V$ *en la clase* $C^1(V')$;

(ii) existen vecindades abiertas Z *de* b *y* Z' *de* 0 *en* $\mathbf{R}^q$, *y una función* $\tau : Z' \to Z$ *en la clase* $C^1(Z')$ *que tiene un inverso* $\tau^{-1} : Z \to Z'$ *en la clase* $C^1(Z)$;

(iii) si $x \in V$ *entonces* $f(x) = \tau \circ i_r \circ \sigma(x)$, *en donde* $i_r : \mathbf{R}^p \to \mathbf{R}^q$ *es la aplicación definida por*

$$i_r(c_1, \ldots, c_r, c_{r+1}, \ldots, c_p) = (c_1, \ldots, c_r, 0, \ldots, 0) \in \mathbf{R}^q.$$

DEMOSTRACION. Se va a suponer que $a = 0$ y $b = 0$ y se usarán la notación y los resultados establecidos en la demostración del teorema de parametrización. Sea $B : \mathbf{R}^p \to \mathbf{R}^p$ la función lineal que transforma a los elementos de la base usual $e_1, \ldots, e_p$ of $\mathbf{R}^p$ a los vectores $x_1, \ldots, x_p$; entonces

B es una biyección de $\boldsymbol{R}^p$ sobre $\boldsymbol{R}^p$ y B^{-1} existe. La aplicación $\sigma:\Omega\to\boldsymbol{R}^p$ definida por $\sigma(x)=B^{-1}\circ u(x)$ pertenece a la clase $C^1(\Omega)$ y puesto que la restricción de u a U tiene un inverso $w:U'\to\boldsymbol{R}^p$ que transforma sobre U, se sigue que la restricción de σ a U tiene un inverso $\sigma^{-1}=w\circ B$ que aplica $B^{-1}(U')$ sobre U.

Sean $W\subseteq\boldsymbol{R}^r$ y $\varphi:W\to\boldsymbol{R}^q$ como en el teorema de parametrización y sea $H:\boldsymbol{R}^q\to\boldsymbol{R}^q$ la función lineal que aplica los elementos de la base usual $e_1,\ldots,e_q$ de $\boldsymbol{R}^q$ a los vectores $y_1,\ldots,y_q$; por lo tanto, H es una biyección de $\boldsymbol{R}^q$ sobre $\boldsymbol{R}^q$ y H^{-1} existe. Se define

$$W'=\{(c_1,\ldots,c_q)\in\boldsymbol{R}^q:(c_1,\ldots,c_r)\in W\}$$

y se considera a $\tau:W'\to\boldsymbol{R}^q$ definido como

$$\tau(c_1,\ldots,c_q)=\varphi(c_1,\ldots,c_r)+H(0,\ldots,0,c_{r+1},\ldots,c_q).$$

Del corolario 41.12(a) se infiere que $D\tau(0)=H$; por lo tanto, el teorema de inversión 41.8 implica que la restricción de τ a alguna vecindad Z' de 0 es una biyección sobre alguna vecindad Z de $\tau(0)=0$.

Restringiendo aún más a V, si es necesario, se puede suponer que $f(V)\subseteq Z$. Ahora, sea $x\in V$ y considere $\sigma(x)=B^{-1}\circ u(x)$. Si i_r se define igual que como se acaba de hacer, entonces $i_r\circ\sigma(x)=(C^{-1}\circ P_1\circ u(x),0)=(\alpha(x),0)$. Por lo tanto, $\tau\circ i_r\circ\sigma(x)=\varphi\circ\sigma(x)=f(x)$ para toda $x\in V$. Q.E.D.

Ejercicios

41.A. Sea $\Omega\subseteq\boldsymbol{R}^p$ abierto y $f:\Omega\to\boldsymbol{R}^q$. Si $Df(x)$ existe para toda $Df(x)$ y si $x\in\Omega$ $i=1,\ldots,q$, $j=1,\ldots,p$, entonces demostrar que $|D_jf_i(x)-D_jf_i(y)|\le\|Df(x)-Df(y)\|_{pq}$. De donde, si f pertenece a la clase $C^1(\Omega)$, entonces cada una de las derivadas parciales D_jf_i es continua en Ω.

41.B. Sea $\Omega\subseteq\boldsymbol{R}^p$ abierto y $f:\Omega\to\boldsymbol{R}^q$. Si f pertenece a la clase $C^1(\Omega)$ y $K\subseteq\Omega$ es compacto, demostrar que $x\mapsto Df(x)$ es uniformemente continua en el sentido de que para toda $\varepsilon>0$ existe $\delta>0$ tal que si $x,y\in K$ y $\|x-y\|<\delta$ entonces $\|Df(x)-Df(y)\|_{pq}<\varepsilon$.

41.C. Sea $\Omega\subseteq\boldsymbol{R}^p$ y $\Omega_1\subseteq\boldsymbol{R}^q$ abiertos y suponga que $f:\Omega\to\boldsymbol{R}^q$ pertenece a la clase $C^1(\Omega)$ y $g:\Omega_1\to\boldsymbol{R}^r$ pertenece a la clase $C^1(\Omega_1)$. Si $f(\Omega)\subseteq\Omega_1$, demostrar que $g\circ f$ pertenec a la clase $C^1(\Omega)$.

41.D. Suponga que $f:\boldsymbol{R}\to\boldsymbol{R}$ está definida por $f(x)=x^3$. Demostrar que f pertenece a la clase $C^1(\boldsymbol{R})$ que es una biyección de $\boldsymbol{R}$ sobre $\boldsymbol{R}$ con inverso $g(x)=x^{1/3}$ para toda $x\in\boldsymbol{R}$. Sin embargo, $Df(0)$ no es ni inyectiva ni suprayectiva. ¿Pertenece g a la clase $C^1(\boldsymbol{R})$?

41.E. Sea $g:\boldsymbol{R}\to\boldsymbol{R}$ tal que $g'(x)\neq0$ para toda $x\in\boldsymbol{R}$. Demostrar que g es una biyección de $\boldsymbol{R}$ sobre $g(\boldsymbol{R})$.

41.F. Sea $A\subseteq\boldsymbol{R}^p$, sea $f:A\to\boldsymbol{R}^p$, y sea $g:f(A)\to\boldsymbol{R}^p$ inverso de f. Supóngase que f es diferenciable en $a\in A$ y que g es diferenciable en $b=f(a)$. Si $Df(a)$ no es invertible, demostrar que $Dg(b)$ no es invertible.

41.G. Suponga que $f:\mathbf{R}^2 \to \mathbf{R}^2$ está dada por

$$f(x, y) = (x + y, 2x + ay).$$

(a) Calcular $Df(x,y)$ y demostrar que $Df(x,y)$ es invertible si y sólo si $a \neq 2$.
(b) Analizar la imagen del cuadrado unitario $\{(x, y): x, y \in [0, 1]\}$ cuando $a = 1, 2, 3$.

41.H. Sea f la transformación de $\mathbf{R}^2$ a $\mathbf{R}^2$ que manda al punto (x, y) en el punto (u, v) y que está dada por

$$u = x, \qquad v = xy.$$

Trazar algunas curvas $u =$ constante, $v =$ constante en el plano (x, y) y algunas curvas $x =$ constante, $y =$ constante en el plano (u, v). ¿Es esta transformación uno a uno? ¿se transforma f sobre todo $\mathbf{R}^2$? Demostrar que si $x \neq 0$, entonces f transforma alguna vecindad de (x, y) de manera uno a uno sobre una vecindad de (x, xy). ¿En qué región en el plano (u, v) transforma f al rectángulo $\{(x, y): 1 \le x \le 2, 0 \le y \le 2\}$? ¿Qué untos en el plano (x, y) se transforman bajo f al rectángulo $\{(u, v): 1 \le u \le 2, 0 \le v \le 2\}$?

41.I. Sea f la transformación de $\mathbf{R}^2$ en $\mathbf{R}^2$ que manda al punto (x, y) en el punto (u, v) y que está dada por

$$u = x^2 - y^2, \qquad v = 2xy.$$

¿Cuáles curvas en el plano (x, y) se transforman bajo f en las rectas $u =$ constante, $v =$ constante? ¿a qué curvas en el plano (u, v) se transforman las rectas $x =$ constante, $y =$ constante? Demostrar que cada punto distinto de cero (u, v) es la imagen bajo f de dos puntos. ¿En qué región aplica f al cuadrado $\{(x, y): 0 \le x \le 1, 0 \le y \le 1\}$? ¿Qué región es transformada por f al cuadrado $\{(u, v): 0 \le u \le 1, 0 \le v \le 1\}$?

41.J. Defina $h:\mathbf{R} \to \mathbf{R}$ como

$$h(x) = x + 2x^2 \operatorname{sen} \frac{1}{x} \qquad \text{para } x \neq 0,$$

$$= 0 \qquad \text{para } x = 0.$$

Demostrar que h no pertenece a la clase $C^1(\mathbf{R})$ y que h no es inyectiva en una vecindad de 0. Sin embargo, es suprayectiva en una vecindad de 0 y $Dh(0)$ es invertible.

41.K. Defina $f:\mathbf{R}^2 \to \mathbf{R}^2$ como $f(x, y) = (y, x + y^2)$ para $(x, y) \in \mathbf{R}^2$. Demostrar que f pertenece a la clase $C^1(\mathbf{R}^2)$ y que f es invertible en alguna vecindad de un punto arbitrario de $\mathbf{R}^2$. Dibujar la imagen bajo f de las rectas $x = 0, \pm 1, \pm 2$ y $y = 0, \pm 1, \pm 2$. Encontrar el inverso $g = f^{-1}:\mathbf{R}^2 \to \mathbf{R}^2$ y demostrar que $Dg(f(x_0, y_0)) = Df(x_0, y_0)^{-1}$.

41.L. (Este ejercicio supone familiarización con el concepto del determinante de una matriz cuadrada.) Sea $L \in \mathscr{L}(\mathbf{R}^p, \mathbf{R}^p)$ y sea $[c_{ij}]$ la representación matricial de L con respecto a la base usual en $\mathbf{R}^p$. En álgebra lineal se demuestra que L es invertible si y sólo si $\Delta = \det[c_{ij}]$ no es cero. Además, si $\Delta \neq 0$, la matriz de L^{-1} es de la forma $[p_{ij}/\Delta]$, en donde las p_{ij} son polinomios en las c_{ij}.

(a) Demostrar que si L_0 es invertible y si $\|L - L_0\|_{pp}$ es suficientemente pequeño, entonces L es invertible.

(b) Demostrar que si L_0 es invertible, entonces la transformación $L \mapsto L^{-1}$ es continua en una vecindad de L_0 con respecto a la norma en $\mathscr{L}(\mathbf{R}^p, \mathbf{R}^p)$.

(c) Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y suponga que $f:\Omega \to \boldsymbol{R}^p$ pertenece a la clase $C^1(\Omega)$. Si $Df(c)$ es invertible par alguna $c \in \Omega$, entonces $Df(x)$ es invertible en alguna vecindad de c.

41.M. Defina $F:\boldsymbol{R}^2 \to \boldsymbol{R}$ como $F(x, y) = y^2 - x$. Demostrar que F pertenece a la clase $C^1(\boldsymbol{R}^2)$ pero que $D_2F(0, 0) = 0$. Demostrar que no existe una función φ definida en una vecindad W de 0 tal que $F(x, \varphi(x)) = 0$ para toda $x \in W$.

41.N. Suponga que $f:\boldsymbol{R}^3 \to \boldsymbol{R}^2$ está definida como

$$f(x, y, z) = (x + y + z, x - y - 2xz),$$

entonces $f(0, 0, 0) = (0, 0)$ y Df(0,0,0,) está dada por

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & -1 & 0 \end{bmatrix}.$$

(a) Demostrar que se puede resolver para $(x, y) = \varphi(z)$ cerca de $z = 0$ y que

$$D\varphi(0) = \begin{bmatrix} -\frac{1}{2} \\ -\frac{1}{2} \end{bmatrix}.$$

(b) Sacar la solución explícita de $(x, y) = \varphi(z)$ para obtener

$$\varphi(z) = \left(\frac{z}{2(z-1)}, \frac{2-2z^2}{2(z-1)}\right) \quad \text{para } z < 1.$$

Verificar el resultado de la parte *(a)*.

(c) Demostrar que se puede resolver para $(y, z) = \psi(x)$ cerca de $x = 0$ y que

$$D\psi(0) = \begin{bmatrix} 1 \\ -2 \end{bmatrix}.$$

(d) Sacar la solución explícita de $(y, z) = \psi(x)$ para obtener

$$\psi(x) = \left(\frac{2x^2 + x}{1 - 2x}, \frac{2x}{2x - 1}\right) \quad \text{para } x < \tfrac{1}{2}.$$

Verificar el resultado de la parte *(c)*.

41.O. Defina $F:\boldsymbol{R}^5 \to \boldsymbol{R}^2$ como

$$F(u, v, w, x, y) = (uy + vx + w + x^2, uvw + x + y + 1),$$

y observe que $F(2, 1, 0, -1, 0) = (0, 0)$.

(a) Demostrar que se puede resolver $F(u, v, w, x, y) = (0, 0)$ para (x, y) en términos de (u, v, w) cerca de (2,1,0).

(b) Si $(x, y) = \varphi(u, v, w)$ es la solución de la parte *(a)*, demostrar que $D\varphi(2, 1, 0)$ está dada por la matriz

$$-\begin{bmatrix} -1 & 2 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 0 & -1 & -3 \\ 0 & 1 & -3 \end{bmatrix}.$$

41.P. Sea $A \subseteq \boldsymbol{R}^3$ y suponga que $F: A \to \boldsymbol{R}$ representa una superficie S_F en $\boldsymbol{R}^3$ implícitamente como la "superficie de nivel"

$$S_F = \{(x, y, z) \in A : F(x, y, z) = 0\}.$$

Si F es diferenciable en un punto $(x_0, y_0, z_0) \in S_F$ que es interior a A, entonces el **espacio tangente** a S_F en este punto es el conjunto de puntos

$$\{(x, y, z) \in \boldsymbol{R}^3 : A_{(x_0, y_0, z_0)}(x, y, z) = 0\},$$

en donde $A_{(x_0, y_0, z_0)}$ es la transformación afín de $\boldsymbol{R}^3 \to \boldsymbol{R}$ definida por

$$\begin{aligned} A_{(x_0, y_0, z_0)}(x, y, z) &= F(x_0, y_0, z_0) + DF(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0) \\ &= DF(x_0, y_0, z_0)(x - x_0, y - y_0, z - z_0). \end{aligned}$$

(a) Demostrar que el espacio tangente en (x_0, y_0, z_0) está dado por

$$\{(x, y, z) : D_1F(x_0, y_0, z_0)(x - x_0) + D_2F(x_0, y_0, z_0)(y - y_0) + D_3F(x_0, y_0, z_0)(z - z_0) = 0\}.$$

Por lo que el espacio tangente a S_F es un plano si al menos uno de los números $D_1F(x_0, y_0, z_0)$, $D_2F(x_0, y_0, z_0)$, $D_3F(x_0, y_0, z_0)$ es diferente de 0. En este caso el espacio tangente a S_F se llama el **plano tangente** S_F en (x_0, y_0, z_0).

41.Q. Suponga que $F : \boldsymbol{R}^3 \to \boldsymbol{R}$, como se da en seguida, representa una superficie S_F en $\boldsymbol{R}^3$ implícitamnte como la superficie de nivel

$$S_F = \{(x, y, z) \in \boldsymbol{R}^3 : F(x, y, z) = 0\}.$$

En cada uno de los siguientes casos determinar el espacio tangente a S_F en los puntos indicados.

(a) Sea $F(x, y, z) = x^2 + y^2 - z$ en los puntos $(1,1,2)$ y $(0,2,4)$.

(b) Sea $F(x, y, z) = x^2 + y^2 + z^2 - 25$ en los puntos $(3,4,0)$ y $(3, 3, \sqrt{7})$.

(c) Sea $F(x, y, z) = z - xy$ en los puntos $(1,1,1)$ y $(4, \frac{1}{2}, 2)$.

41.R. *(a)* Suponga que además de las hipótesis el teorema de inversión 41.8 se sabe que la función f tiene derivadas parciales continuas de orden $m > 1$. Demostrar que la función inversa $g : V \to \boldsymbol{R}^p$ tiene derivadas parciales continuas de orden m.

(b) Demostrar el resultado análogo para el teorema de la función implícita 41.9.

41.S. Suponga que $f : \boldsymbol{R}^2 \to \boldsymbol{R}$ pertenece a la clase $C^1(\boldsymbol{R}^2)$. Demostar que f no es inyectiva; de hecho, la restricción de f a cualquier conjunto abierto de $\boldsymbol{R}^2$ no es inyectiva.

41.T. Suponga que $g : \boldsymbol{R} \to \boldsymbol{R}^2$ pertenece a la clase $C^1(\boldsymbol{R})$. Demostrar que si $c \in \boldsymbol{R}$, entonces la restricción de g a cualquier vecindad de c no es una transformación suprayectiva sobre una vecindad de $g(c)$.

41.U. Sea $L \in \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ inyectiva y sea $r > 0$ tal que $r\,\|x\| \leq \|L(x)\|$ para toda $x \in \boldsymbol{R}^p$. Demostrar que si $L_1 \in \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ es tal que $\|L_1 - L\|_{pq} < r$, entonces L_1 es inyectiva. (Por lo tanto, el conjunto de transformaciones inyectivas es abierto en $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$.)

41.V. Sea $L \in \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ suprayectiva y s a $m > 0$ como en la demostración de 41.6 Demostrar que si $L_1 \in \mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$ s tal que $\|L_1 - L\|_{pq} < m/2$, entonces L_1 es suprayectiva. (Por lo tanto, el conjunto de transformaciones suprayectivas es abierto en $\mathscr{L}(\boldsymbol{R}^p, \boldsymbol{R}^q)$.

41.W. Suponga que $g : \boldsymbol{R}^p \to \boldsymbol{R}^p$ pertenece a la clase $C^1(\boldsymbol{R}^p)$ y satisface $\|Dg(x)\|_{pp} \leq a < 1$ para toda $x \in \boldsymbol{R}^p$. Si $f(x) = x + g(x)$ para $x \in \boldsymbol{R}^p$, demostrar que f satisface

$$\|f(x_1) - f(x_2) - (x_1 - x_2)\| \leq a\,\|x_1 - x_2\|$$

para toda x_1, x_2 en $\boldsymbol{R}^p$ y que f es una biyección de $\boldsymbol{R}^p$ sobre $\boldsymbol{R}^p$.

Proyectos

41. α. (Este proyecto da una demostración directa y elemental del teorema de la función implícita.) Sea $\Omega \subseteq \mathbf{R}^2$ abierto y suponga que $F:\Omega \to \mathbf{R}$ pertenece a la clase $C^1(\Omega)$. Suponga que $(a, b) \in \Omega$, que $F(a, b) = 0$, y que $D_2F(a, b) > 0$.

(a) Demostrar que existe una celda cerrada $Q = [a_1, a_2] \times [b_1, b_2]$ n centro (a, b) tal que $D_2F(x, y) > 0$ para toda $(x, y) \in Q$, y tal que $F(x, b_1) < 0$ y $F(x, b_2) > 0$ para toda $x \in [a_1, a_2]$.

(b) Si $x \in [a_1, a_2]$, entonces la función $F_x:[b_1, b_2] \to \mathbf{R}$ definida como $F_x(y) = F(x, y)$ para $y \in [b_1, b_2]$ es tal que $F_x(b_1) < 0 < F_x(b_2)$ y $F_x'(y) > 0$ para $y \in [b_1, b_2]$.

(c) Existe una función φ que transforma a $[a_1, a_2]$ en $[b_1, b_2]$ tal que $F(x, \varphi(x)) = 0$ para toda $x \in [a_1, a_2]$.

(d) Si $x \in (a_1, a_2)$ y $|h|$ es suficientemente pequeña, demostrar que existe h_1 con $0 < |h_1| < |h|$ tal que

$$\begin{aligned} 0 &= F[x + h, \varphi(x + h)] - F[x, \varphi(x)] \\ &= D_1F[x + h_1, \varphi(x + h_1)]h + D_2F[x + h_1, \varphi(x + h_1)][\varphi(x + h) - \varphi(x)]. \end{aligned}$$

(e) Demostrar que φ es diferenciable en (a_1, a_2) y que $\varphi'(x) = -D_1F[x, \varphi(x)]/D_2F[x, \varphi(x)]$.

(f) Modificar el argumento anterior para una función F definida en unconjunto abierto $\Omega \subseteq \mathbf{R}^p$.

(g) Sea $\Omega \subseteq \mathbf{R}^p \times \mathbf{R}^2$ abierto y suponga que $F, G:\Omega \to \mathbf{R}$ pertenecen a la clase $C^1(\Omega)$ y que para algún punto $(a, b) \in \mathbf{R}^p \times \mathbf{R}^2$ se tiene $F(a, b) = 0$, $G(a, b) = 0$. Suponga que

$$\Delta = \det \begin{bmatrix} D_{p+1}F(a, b) & D_{p+2}F(a, b) \\ D_{p+1}G(a, b) & D_{p+2}G(a, b) \end{bmatrix} \neq 0,$$

entonces al menos una de las dos $D_{p+1}F(a, b)$ y $D_{p+2}F(a, b)$ es distinta de cero. Suponga que $D_{p+2}F(a, b) \neq 0$ y usar *(f)* para obtener $x_{p+2} = \varphi(x_1, \ldots, x_{p+1})$ en una vecindad de $(a, b_1) \in \mathbf{R}^p \times \mathbf{R}$. Por lo que $F(x_1, \ldots, x_{p+1}, \varphi(x_1, \ldots, x_{p+1})) = 0$, en esta vecindad. Sea ahora

$$H(x_1, \ldots, x_{p+1}) = G(x_1, \ldots, x_{p+1}, \varphi(x_1, \ldots, x_{p+1})).$$

Por la regla de la cadena

$$D_{p+1}H = D_{p+1}G + (D_{p+2}G)(D_{p+1}\varphi),$$

en donde estas funciones se calculan en los puntos apropiados. Dado que $D_{p+1}\varphi = -(D_{p+1}F)/(D_{p+2}F)$ deduce que $D_{p+1}H = -\Delta/D_{p+2}F$ que no es cero en (a, b_1). Por lo tanto, se puede usar *(f)* para obtener $x_{p+1} = \psi(x_1, \ldots, x_p)$ en una vecindad de $a \in \mathbf{R}^p$. (Esto prueba el teorema de la función implícita en el caso en que $q = 2$; por inducción se obtienen extensiones para el caso general de q.)

41. β. (Este proyecto es paralelo al proyecto 40.α da una demostración más directa de la primera parte del teorema de inversión 41.8 que la que se da el texto) Se va a suponer que $\Omega \subseteq \mathbf{R}^p$ es abieto, que $f:\Omega \to \mathbf{R}^p$ pertenece a la clase $C^1(\Omega)$, y que para alguna $x_0 \in \Omega$, la transformación lineal $Df(x_0)$ es una biyección. Sea $\Gamma = Df(x_0)^{-1}$.

(a) Demostrar que existe $r > 0$ tal que si $\|x - x_0\| \leq r$, entonces $\|I - \Gamma \circ Df(x)\|_{pp} \leq \frac{1}{2}$.

(b) Sea $s \leq \frac{1}{2} r \|\Gamma\|_{pp}^{-1}$, para y fija con $\|y - f(x_0)\| \leq s$, se define $F_y(x) = f(x) - y$ para

$\|x - x_0\| \leq r$. entonces, F_y es diferenciable $\|\Gamma \circ F_y(x_0)\| \leq \frac{1}{2}r$, y $\|I - \Gamma \circ DF_y(x)\|_{pp} \leq \frac{1}{2}$ para $\|x - x_0\| \leq r$.

(c) Si $\|y - f(x_0)\| \leq s$, defina G_y para $\|x - x_0\| \leq r$ como $G_y(x) = x - \Gamma \circ F_y(x)$. Entonces, G_y es una contracción con constante $\frac{1}{2}$ en esta bola.

(d) Si $\|y - f(x_0)\| \leq s$, definir $\varphi(y) = x_0$ y $\varphi_{n+1}(y) = G_y(\varphi_n(y))$ para $n = 0, 1, 2, \ldots$. Demostrar que $\|\varphi_{n+1}(y) - \varphi_n(y)\| \leq 2^{-n} \|\varphi_1(y) - \varphi_0(y)\| \leq 2^{-n-1}r$, por lo que se sigue que $\|\varphi_{n+1}(y) - \varphi_m(y)\| \leq 2^{-m}r$ para $n \geq m \geq 0$. En particular, $\|\varphi_k(y) - x_0\| \leq r$, de manera que es posible realizar esta iteración.

(e) Demostrar que cada una de las funciones φ_k es para $\|y - f(x_0)\| \leq s$ y que la sucesión (φ_k) es uniformemente convergente a una función continua φ que es tal que $G_y(\varphi(y)) = \varphi(y)$ para $\|y - f(x_0)\| \leq s$, por lo que se infiere que $f(\varphi(y)) = y$ para $\|y - f(x_0)\| \leq s$. Por lo tanto, la función φ es el inverso de f en el conjunto $\{y : \|y - f(x_0)\| \leq s\}$ y lo transforma en el conjunto $\{x : \|x - x_0\| \leq r\}$.

41. γ. (Este proyecto es paralelo a los proyectos 40.α y 41.β y da una demostración directa del teorema de la función implícita.) Sea $\Omega \subseteq \boldsymbol{R}^p \times \boldsymbol{R}^q$ abierto y sea $(x_0, y_0) \in \Omega$. Suponga que $F : \Omega \to \boldsymbol{R}^q$ pertenece a la clase $C^1(\Omega)$, que $F(x_0, y_0) = 0$, y que la transformación lineal $L_2 : \boldsymbol{R}^q \to \boldsymbol{R}^q$ definida por

$$L_2(v) = DF(x_0, y_0)(0, v) \qquad \text{para } v \in \boldsymbol{R}^q$$

es una biyección de $\boldsymbol{R}^q$ sobre $\boldsymbol{R}^q$. Sea $\Gamma = L_2^{-1}$.

(a) Demostrar que existe $r > 0$ tal que si $\|x - x_0\|^2 + \|y - y_0\|^2 \leq r^2$, entonces

$$\|v - \Gamma \circ DF(x, y)(0, v)\| \leq \tfrac{1}{2}\|v\| \qquad \text{para } v \in \boldsymbol{R}^q.$$

(b) Sea $0 < s \leq \frac{1}{2}r$ tal que si $\|x - x_0\| \leq s$, entonces

$$\|F(x, y_0)\| \leq \tfrac{1}{4}r\,\|\Gamma\|_{qq}^{-1}.$$

Para cada x fija con $\|x - x_0\| \leq s$, se define $G_x(y) = y - \Gamma \circ F(x, y)$ para $\|y - y_0\| \leq \frac{1}{2}r$, con valores en $\boldsymbol{R}^q$. Demostrar que para cada x con $\|x - x_0\| \leq s$, se tiene

$$\|G_x(y_1) - G_x(y_2)\| \leq \tfrac{1}{2}\|y_1 - y_2\|$$

para todas y_1, y_2 que satisfagan $\|y_j - y_0\| \leq \frac{1}{2}r$.

(c) Si $\|x - x_0\| \leq s$, defina $\psi_0(x) = y_0$ y $\psi_{n+1}(x) = G_x(\psi_n(x))$ para $n = 0, 1, 2, \ldots$. Demostrar que $\|\psi_{n+1}(x) - \psi_m(x)\| \leq 2^{-m-1}r$ para $n \geq m \geq 0$. Entonces, $\|\psi_k(x) - y_0\| \leq \frac{1}{2}r$, por lo que la iteración es posible.

(d) Demostrar que cada una de las funciones ψ_k es continua para $\|x - x_0\| \leq s$ y que la sucesión(ψ_k)es uniformemente convergente a una función continua ψ tal que

$$F(x, \psi(x)) = 0 \qquad \text{for all } \|x - x_0\| \leq s.$$

(e) Para probar que ψ es diferenciable para $\|x = x_0\| < s$ usar el ejercicio 39.W y emplear un argumento análogo al de los incisos *(d)* y *(e)* del proyecto 41. α para cada componente de F.

Sección 42 Problemas sobre extremo

En la sección 27 se analizó brevemente el conocido proceso para localizar puntos interiores en los que una función diferenciable de valor real de una variable adquiere valores extremos relativos. La pregunta de si un **punto crí-**

tico (es decir, un punto en que la derivada es cero) es realmente un punto extremo no siempre se analiza pero a menudo se puede manejar por medio del teorema de Taylor 28.6. El análisis de puntos extremos que pertenecen a la frontera del dominio con frecuencia se somete a una aplicación del teorema del valor medio 27.6.

En el caso de una función con dominio en $\mathbf{R}^p$ $(p > 1)$ y rango en $\mathbf{R}$, la situación es a veces más complicada y es necesario estudiar cada función en forma independiente. Sin embargo, existen algunos teoremas generales que son útiles y se presentarán en seguida.

Sea $\Omega \subseteq \mathbf{R}^p$ y sea $f: \Omega \to \mathbf{R}$. Un punto $c \in \Omega$ se dice que es un punto de mínimo relativo de f si existe $\delta > 0$ tal que $f(c) \leq f(x)$ para toda $x \in \Omega$ con $\|x - c\| < \delta$. Un punto $c \in \Omega$ se dice que es un **punto de mínimo relativo estricto** de f si existe $\delta > 0$ tal que $f(c) < f(x)$ para toda $x \in \Omega$ con $0 < \|x - c\| < \delta$. Se define un **punto de máximo relativo (estricto)** de f análogamente. Más aún, si $c \in \Omega$ es un punto de mínimo relativo (estricto) de f, se dice que c es un **punto extremo relativo (estricto)** de f o que f tiene un **extremo relativo (estricto)** en c.

Con frecuencia es útil el siguiente resultado.

42.1 TEOREMA. *Sea* $\Omega \subseteq \mathbf{R}^p$, *y sea* $f: \Omega \to \mathbf{R}$. *Si un punto interior de* c *es un punto de extremo relativo de* f *y si la derivada parcial* $D_u f(c)$ *de* f *con respecto a un vector* $u \in \mathbf{R}^p$ *existe, entonces* $D_u f(c) = 0$.

DEMOSTRACION. Por hipótesis, la restricción de f a la intersección de Ω con la recta $\{c + tu : t \in \mathbf{R}\}$ tiene un extremo relativo en c. Por lo tanto, del teorema 27.4 se sigue que $D_u f(c) = 0$. Q.E.D.

42.2 COROLARIO. *Sea* $\Omega \subseteq \mathbf{R}^p$, *y sea* $f: \Omega \to \mathbf{R}$. *Si un punto interior* c *de* Ω *es un punto de extremo relativo de* f *y si la derivada* $Df(c)$ *existe, entonces* $Df(c) = 0$.

DEMOSTRACION. Del corolario 39.7 se deduce que cada una de las derivadas parciales $D_j f(c)$, $j = 1, \ldots, p$, existe y que si $u = (u_1, \ldots, u_p) \in \mathbf{R}^p$, entonces

$$Df(c)(u) = \sum_{j=1}^{p} u_j D_j f(c).$$

Por el teorema anterior, $D_j f(c) = 0$ para $j = 1, \ldots, p$, por lo que $Df(c)(u) = 0$ para toda $u \in \mathbf{R}^p$. Q.E.D.

Se sigue que si $\Omega \subseteq \mathbf{R}^p$, y si $f: \Omega \to \mathbf{R}$ tiene un extremo relativo en $c \in \Omega$ y $Df(c)$ existe, entonces

$$(42.1) \qquad D_1 f(c) = 0, \ldots, D_p f(c) = 0.$$

Un punto interior c en el que $Df(c) = 0$ se llama **punto crítico** de f. Se deduce que si Ω es un conjunto abierto en $\mathbf{R}^p$ en donde f es diferenciable, entonces el

conjunto de puntos críticos de f contendrá a todos los puntos extremos relativos de f. Desde luego, este conjunto de puntos críticos también puede contener puntos en los que f no tenga un extremo relativo. (Además, la función f puede tener un extremo relativo en un punto interior c de Ω en el que la derivada $Df(c)$ no existe o f puede tener un extremo relativo en un punto $c \in \Omega$ que no es punto interior de Ω; en cualquiera de los casos, el punto c no será un punto crítico de f.)

42.3 EJEMPLOS. *(a)* Sea $f_1(x) = x^3$ para $x \in [-1, 1]$. Entonces, $Df_1(0) = 0$; sin embargo, f_1 no tiene un extremo en $x = 0$. Por otro lado, f_1 tiene extremos estrictos en los puntos ± 1 (que no son puntos interiores del dominio y no son puntos críticos).

(b) Sea $f_2(x) = |x|$ para $x \in [-1, 1]$. Entonces, $Df_2(0)$ no existe; sin embargo, f_2 tiene un mínimo relativo estricto en el punto interior 0. Por otro lado, f_2 tiene extremos relativos estrictos en los puntos ± 1.

(c) Suponga que $f_3: \mathbf{R}^2 \to \mathbf{R}$ está definida por $f_3(x, y) = xy$. Entonces, $Df_3(0, 0) = 0$ de manera que el origen (0,0) es un punto crítico de f_3; sin embargo, no es un extremo relativo de f_3 ya que

$$f_3(0, 0) < f_3(x, y) \qquad \text{para } xy > 0,$$
$$f_3(0, 0) > f_3(x, y) \qquad \text{para } xy < 0.$$

Se dice que el origen (0,0) es un **punto silla** de f_3 esto quiere decir que toda vecindad de (0,0) contiene puntos en los que f_3 es estrictamente mayor que $f_3(0, 0)$ y también contiene puntos en los que f_3 es estrictamente menor que $f_3(0, 0)$.

(d) Suponga que $f_4: \mathbf{R}^2 \to \mathbf{R}$ está definida por $f_4(x, y) = (y - x^2)(y - 2x^2)$. Demostrar que $Df_4(0, 0) = 0$ y que la restricción de f_4 a toda recta que pasa por (0,0) tiene un mínimo relativo en el origen. Sin embargo, demostrar que en toda vecindad de (0,0) existen puntos en los que f_4 es estrictamente positiva y otros en los que f_4 es estrictamente negativa.

Prueba de la segunda derivada

En vista de los ejemplos que se acaban de dar, es conveniente tener condiciones que sean necesarias (o sean suficientes) para garantizar que un punto crítico es un extremo o es un punto silla. Los siguientes resultados dan condiciones en términos de la segunda derivada de f que se introdujo al final de la sección 40.

42.4 TEOREMA. *Sea $\Omega \subseteq \mathbf{R}^p$ abierto y suponga que $f: \Omega \to \mathbf{R}$ tiene segundas derivadas parciales continuas en Ω. Si $c \in \Omega$ es un punto de mínimo relativo [respectivamente, máximo] de f, entonces*

$$D^2 f(c)(w)^2 = \sum_{i, j=1}^{p} D_{ij} f(c) w_i w_j \geq 0 \tag{42.2}$$

[respectivamente, $D^2 f(c)(w)^2 \leq 0$] para toda $w \in \mathbf{R}^p$.

DEMOSTRACION. Sea $w \in \mathbf{R}^p$, $\|w\| = 1$. Si c un punto de mínimo relativo, existe $\delta > 0$ tal que si $|t| < \delta$; entonces, $f(c + tw) - f(c) \geq 0$. Dado que Ω es abierto, existen $\delta_1 > 0$ con $\delta_1 \leq \delta$ tal que $c + tw$ pertenece a Ω para $0 \leq t \leq \delta_1$. Por el teorema de Taylor 40.9 existe t_1 con $0 \leq t_1 \leq t \leq \delta_1$ tal que si $c_t = c + t_1 w$, entonces,

$$f(c + tw) = f(c) + Df(c) + \tfrac{1}{2}D^2 f(c_t)(tw)^2.$$

Dado que c es un punto de mínimo relativo, del corolario 42.2 se infiere que $Df(c) = 0$; por lo tanto, se tiene

$$\tfrac{1}{2}D^2 f(c_t)(tw)^2 \geq 0$$

para $0 \leq t \leq \delta_1$. Se sigue que $D^2 f(c_t)(w)^2 \geq 0$. Dado que $\|c_t - c\| = |t_1| \leq |t|$, se deduce que $c_t \to c$ conforme $t \to 0$. Como las segundas derivadas parciales de f son continuas, entonces, $D^2 f(c)(w)^2 \geq 0$ para toda $w \in \mathbf{R}^p$ con $\|w\| = 1$, de donde se obtiene el resultado. Q.E.D.

El siguiente resultado es un inverso parcial del teorema 42.4. Sin embargo, observe que su hipótesis es un poco más fuerte que la conclusión en 42.4

42.5 TEOREMA. *Sea $\Omega \subseteq \mathbf{R}^p$ abierto, suponga que $f: \Omega \to \mathbf{R}$ tiene segundas derivadas parciales continuas en Ω, y sea $c \in \Omega$ un punto crítico de f.*

(a) Si $D^2 f(c)(w)^2 > 0$ para toda $w \in \mathbf{R}^p$, $w \neq 0$, entonces f tiene un mínimo relativo estricto en c.

(b) Si $D^2 f(c)(w)^2 < 0$ para toda $w \in \mathbf{R}^p$, $w \neq 0$, entonces f tiene un máximo relativo estricto en c.

(c) Si $D^2 f(c)(w)^2$ toma valores estrictamente positivos así como estrictamente negativos para $w \in \mathbf{R}^p$, entonces f tiene un punto silla en c.

DEMOSTRACION. *(a)* Por hipótesis, $D^2 f(c)(w)^2 > 0$ para w en el conjunto compacto $\{w \in \mathbf{R}^p : \|w\| = 1\}$. Dado que la aplicación $w \mapsto D^2 f(c)(w)^2$ es continua, existe $m > 0$ tal que

$$D^2 f(c)(w)^2 \geq m \quad \text{para } \|w\| = 1.$$

Dado que las segundas derivadas parciales de f son continuas en Ω, existe $\delta > 0$ tal que si $\|x - c\| < \delta$ entonces

$$D^2 f(x)(w)^2 \geq \tfrac{1}{2} m \quad \text{para } \|w\| = 1.$$

Por el teorema del Taylor 40.9 si $0 \leq t \leq 1$ hay un punto c_t en el segmento de línea que une a c con $c + tw$ tal que

$$f(c + tw) = f(c) + Df(c)(tw) + \tfrac{1}{2}D^2 f(c_t)(tw)^2.$$

Dado que c es un punto crítico, se infiere que si $\|w\| = 1$ y $0 < t < \delta$, entonces

$$f(c + tw) - f(c) = \tfrac{1}{2}t^2 D^2 f(c_t)(w)^2 \geq \tfrac{1}{4}mt^2 > 0,$$

De modo que $f(c+u)>f(c)$ para $0<\|u-c\|<\delta$, y f tiene un mínimo relativo estricto en c. Por lo que el inciso *(a)* está demostrado y la demostración del inciso *(b)* es análoga.

(c) Sean w_+, w_- vectores unitarios en $\boldsymbol{R}^p$ tales que

$$D^2f(c)(w_+)^2>0, \qquad D^2f(c)(w_-)^2<0.$$

Del teorema de Taylor se deduce que para $t>0$ suficientemente pequeña se tiene

$$f(c+tw_+)>f(c), \qquad f(c+tw_-)<f(c).$$

Por lo tanto, c es un punto silla de f.

Comparando los teoremas 42.4 y 42.5, se tiende a hacer las siguientes conjeturas: *(i)* si $c\in\Omega$ es un punto de mínimo relativo estricto, entonces $D^2f(c)(w)^2>0$ para toda $w\in\boldsymbol{R}^p$, $w\neq 0$, *(ii)* si $c\in\Omega$ es un punto silla de f, entonces $D^2f(c)(w)^2$ toma valores estrictamente positivos, así como estrictamente negativos *(iii)* si $D^2f(c)(w)^2\geq 0$ para toda $w\in\boldsymbol{R}^p$, entonces c es un punto de mínimo relativo. Todas estas suposiciones son falsas como se puede ver con ejemplos.

Para poder implementar el teorema 42.5 es necesario saber si la función $w\mapsto D^2f(c)(w)^2$ es de un signo. Un resultado importante y muy conocido del álgebra (véase el libro de Hoffman y Kunze citado en las referencias) se puede usar para determinar esto. Para cada $j=1,2,\ldots,p$, sea Δ_j el determinante de la matriz (simétrica)

$$\begin{bmatrix} D_{12}f(c) & \cdots & D_{1j}f(c) \\ \vdots & & \vdots \\ D_{j1}f(c) & \cdots & D_{jj}f(c) \end{bmatrix}.$$

Si los números $\Delta_1,\Delta_2,\ldots,\Delta_p$ son todos estrictamente positivos, entonces $D^2f(c)(w)^2>0$ para toda $w\neq 0$ y f tiene un mínimo relativo estricto en c. Si los números $\Delta_1,\Delta_2,\ldots,\Delta_p$ son en forma alternante estrictamente negativos y estrictamente positivos, entonces $D^2f(c)(w)^2<0$ para toda $w\neq 0$ y f tiene un máximo relativo estricto en c. En otros casos puede haber puntos extremos o puntos silla.

En el caso especial e importante $p=2$, es más conveniente una formulación menos elaborada y se puede obtener un poco más de información. En este caso es necesario analizar la función cuadrática

$$Q=Au^2+2Buv+Cv^2.$$

Si $\Delta=AC-B^2>0$, entonces $A\neq 0$ (y $C\neq 0$) y se puede completar el cuadrado y escribir

$$Q=\frac{1}{A}[(Au+Bv)^2+(AC-B^2)v^2].$$

De modo que el signo de Q es el mismo que el de A(o C). Por otro lado, si $\Delta<0$, entonces Q tiene valores estrictamente positivos, así como estrictamente negativos. Esto es obvio por la ecuación anterior si $A\neq 0$ y también se demuestra con facilidad si $A=0$.

Estas observaciones se juntan en una proposición formal.

42.6 COROLARIO. *Sea* $\Omega\subseteq \boldsymbol{R}^2$ *abierto, suponga que* $f:\Omega\to\boldsymbol{R}$ *tiene segundas derivadas parciales continuas en* Ω, *sea* $c\in\Omega$ *un punto crítico de* f *y sea*

$$(42.3)\qquad \Delta = D_{11}f(c)D_{22}f(c)-[D_{12}f(c)]^2.$$

(a) Si $\Delta>0$ *y si* $D_{11}f(c)>0$, *entonces* f *tiene un mínimo relativo estricto en* c.

(b) Si $\Delta>0$ *y si* $D_{11}f(c)<0$, *entonces* f *tiene un máximo relativo estricto en* c.

(c) Si $\Delta<0$, *entonces tiene un punto silla en* c.

En los ejercicios se dará cierta información concerniente al caso en que $\Delta=0$

Problemas de extremos con restricciones

Hasta este momento se ha estado analizando el caso en que los extremos de la función $f:\Omega\to\boldsymbol{R}$ pertenecen al interior de su dominio $\Omega\subseteq\boldsymbol{R}^p$. Ninguna de las observaciones es aplicable a la localización de los extremos en la frontera. Sin embargo, si la función está definida en la frontera de Ω y si esta frontera de Ω se puede parametrizar por una función φ, entonces el problema del extremo se reduce al análisis de los extremos de la composición $f\circ\varphi$.

Hay un problema al respecto que origina un procedimiento interesante y elegante. Suponga que S es una "superficie" contenida en el dominio Ω de la función de valor real f. A menudo es deseable encontrar los valores de f que sean máximo o mínimo entre todos los que se tienen en S. Por ejemplo, si $\Omega=\boldsymbol{R}^p$ y $f(x)=\|x\|$, entonces en el problema que se ha planteado se deben encontrar los puntos en la superficie S que estén más cerca o más lejos del origen. Si la superficie S se da paramétricamente, entonces este problema se puede tratar considerando la composición de f con la representación paramétrica de S. Sin embargo, con frecuencia no es conveniente expresar S de este modo y a menudo es deseable otro procedimiento.

Suponga que S se puede dar como los puntos x en Ω que satisfacen una relación de la forma.

$$g(x)=0,$$

para una función g definida de Ω a $\boldsymbol{R}$. Se está intentando encontrar los valores extremos relativos de f para aquellos puntos x en Ω que satisfagan la **restricción** (o **condición lateral**) $g(x)=0$. Si se supone que f y g están en la

clase $C^1(\Omega)$ y que $Dg(c) \neq 0$, entonces una condición necesaria para que c sea un punto extremo de f con respecto a los puntos x que satisfacen $g(x) = 0$, es que la derivada $Dg(c)$ sea un múltiplo de $Df(c)$. En términos de derivadas parciales, esta condición significa que existe un número real λ tal que

(42.4)
$$\begin{aligned} D_1 f(c) &= \lambda D_1 g(c), \\ &\cdots\cdots\cdots \\ D_p f(c) &= \lambda D_p g(c). \end{aligned}$$

En la práctica se desea determinar las p coordenadas del punto c que satisfagan esta condición necesaria. Sin embargo, tampoco se conoce el número real λ, al que por lo general se le llama el **multiplicador de Lagrange.** Las p ecuaciones dadas antes junto con la ecuación

$$g(c) = 0$$

se resuelven para las $p+1$ cantidades desconocidas de las cuales las coordenadas de c son de interés primordial.

42.7 TEOREMA DE LAGRANGE. *Sea $\Omega \subseteq \mathbf{R}^p$ abierto y suponga que f y g son funciones de valor real en la clase $C^1(\Omega)$. Suponga que $c \in \Omega$ es tal que $g(c) = 0$ y que existe una vecindad U de C tal que*

$$f(x) \leq f(c) \qquad [\text{ó } f(x) \geq f(c)]$$

para todos los puntos $x \in U$ que satisfacen $g(x) = 0$. Entonces, existen números reales μ, λ, no siendo igual a cero ambos, tales que

(42.5)
$$\mu Df(c) = \lambda Dg(c).$$

Más aún, si $Dg(c) \neq 0$, se puede tomar $\mu = 1$.

DEMOSTRACION. Suponga que $F: U \to \mathbf{R}^2$ está definida por

$$F(x) = (f(x), g(x)) \qquad \text{para } x \in U,$$

de tal manera que F pertenece a la clase $C^1(U)$ y

$$DF(x)(v) = (Df(x)(v), Dg(x)(v)), \qquad x \in U, \ v \in \mathbf{R}^p,$$

Además, un punto $x \in U$ satisface la restricción $g(x) = 0$ si y sólo si $F(x) = (f(x), 0)$.

Si $f(x) \leq f(c)$ para toda $x \in U$ que satisfaga $g(x) = 0$, entonces los puntos $(r, 0)$ con $f(c) < r$ no están en la imagen $F(U)$; por lo tanto, $DF(c)$ no es una suprayección de $\mathbf{R}^p$ sobre $\mathbf{R}^2$. Pero dado que, la imagen de la transformación lineal $DF(c)$ está contenida en alguna recta en $\mathbf{R}^2$ que pasa por (0,0). Por lo tanto, existe un punto $(\lambda, \mu) \neq (0, 0)$ tal que la imagen de $DF(c)$ está contenida en la recta que pasa por (0,0) y (λ, μ). Por lo que se tiene

(42.6)
$$\mu Df(c)(v) = \lambda Dg(c)(v) \quad \text{para toda } v \in \mathbf{R}^p, \ v \in \mathbf{R}^p.$$

y la ecuación (42.5) se sigue.

Por último, suponga que $Dg(c) \neq 0$. Si $\mu = 0$, la ecuación (42.5) implica que $\lambda = 0$, lo que contradice el hecho de que $(\mu, \lambda) \neq (0, 0)$. Por lo tanto, en este caso se debe tener $\mu \neq 0$ y así poder dividir entre μ y reemplazar λ/μ por λ. Q.E.D.

Dado que $U \subseteq \boldsymbol{R}^p$, la ecuación (42.6) con $v = e_1, \ldots, e_p$ origina el sistema de p ecuaciones:

$$\mu D_1 f(c) = \lambda D_1 g(c),$$
$$\cdots\cdots\cdots\cdots$$
$$\mu D_p f(c) = \lambda D_p g(c).$$

Si no desaparecen todas las $D_i g(c)$, $i = 1, \ldots, p$, entonces se puede tomar $\mu = 1$ para obtener el sistema (42.4)

Se debe hacer hincapié que el teorema de Lagrange da sólo una condición necesaria y que los puntos que se obtienen al resolver las ecuaciones (que a menudo es difícil hacer) pueden ser máximos relativos, mínimos relativos o ninguno de los dos. Sin embargo, con frecuencia se puede usar el corolario 41.13 para probar máximos o mínimos relativos. Más aún, en muchas aplicaciones la determinación de si los puntos son realmente extremos se puede basar en consideraciones geométricas o físicas.

42.8 EJEMPLOS. *(a)* Se desea encontrar un punto en el plano $\{(x, y, z): 2x + 3y - z = 5\}$ en R^3 que sea lo más cercano al origen. Para resolver este problema se va a minimizar la función que da el cuadrado de la distancia al origen:

$$f(x, y, z) = x^2 + y^2 + z^2,$$

con la restricción

$$g(x, y, z) = 2x + 3y - z - 5 = 0.$$

Dado que $Dg(c) \neq 0$ para toda $c \in R^3$, el teorema de Lagrange nos lleva al sistema

$$2x = 2\lambda,$$
$$2y = 3\lambda,$$
$$2z = -\lambda,$$
$$2x + 3y - Z - 5 = 0.$$

Por lo tanto, al subsituir x, y, z se obtiene

$$2\lambda + 3(\tfrac{3}{2}\lambda) - (-\tfrac{1}{2}\lambda) - 5 = 0,$$

o $14\lambda = 4\lambda + 9\lambda + \lambda = 10$, por lo que $\lambda = 5/7$. y se llega al punto $(5/7,\ 15/14,\ -5/14)$. en el plano más cercano a $(0,0,0)$.

(b) Encontrar las dimensiones de la caja rectangular, abierta por arriba, con volumen máximo y área de la superficie dada A. Sean x, y, z las dimensiones de la caja con z como altura. Entonces, se desea maximizar la función

$$V(x, y, z) = xyz$$

sujeta a la restricción

$$g(x, y, z) = xy + 2xz + 2yz - A = 0.$$

Dado que el punto deseado tendrá coordenadas estrictamente positivas, el teorema de Lagrange da origen al sistema

$$yz = \lambda(y + 2z),$$
$$xz = \lambda(x + 2z),$$
$$xy = \lambda(2x + 2y),$$
$$xy + 2xz + 2yz - A = 0.$$

Si se multiplican las tres primeras ecuaciones por x, y, y z, respectivamente, se igualan y se dividen por λ (¿por qué es $\lambda \neq 0$?), se llega a

$$xy + 2xz = xy + 2yz = 2xz + 2yz.$$

La primera igualdad implica $x = y$, y la segunda implica $y = 2z$. Por lo tanto, las razones de los lados son 2:2:1 y de la última ecuación se infiere que $4z^2 + 4z^2 + 4z^2 = A$ lo que implica que $z = \frac{1}{2}(A/3)^{1/2}$. Por lo tanto, el volumen de esta caja es $\frac{1}{2}(A/3)^{3/2}$.

Con frecuencia hay más de una restricción; en este caso, el siguiente resultado es útil.

42.9 TEOREMA. *Sea $\Omega \subseteq \mathbf{R}^p$ abierto y suponga que f y $g_1, \ldots, g_k$ son funciones de valor real en $C^1(\Omega)$. Suponga que $c \in \Omega$ satisface las restricciones*

$$g_1(x) = 0, \ldots, g_k(x) = 0,$$

y que existe una vecindad abierta U de a tal que $f(x) \leq f(c)$ [o $f(x) \geq f(c)$] para toda $x \in U$ que satisfaga estas restricciones. Entonces, existen números reales μ, $\lambda_1, \ldots, \lambda_k$ no todos igual a cero tales que

$$\mu Df(c) = \lambda_1 Dg_1(c) + \cdots + kDg_k(c). \tag{42.7}$$

DEMOSTRACION. Suponga que $F: U \to \mathbf{R}^{k+1}$ está definida por

$$F(x) = (f(x), g_1(x), \ldots, g_k(x)) \qquad \text{for } x \in U,$$

y use el mismo argumento que en la demostración del teorema 42.7.

Q.E.D.

42.10 COROLARIO. *Además de las hipótesis del teorema* 42.9, *suponga que el rango de la matriz*

$$\begin{bmatrix} D_1g_1(c) & \cdots & D_1g_k(c) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ D_pg_1(c) & \cdots & D_pg_k(c) \end{bmatrix} \tag{42.8}$$

es igual a $k(\leq p)$. *Entonces, existen números reales* $\lambda_1, \ldots, \lambda_k$ *no todos iguales a cero tales que*

(42.9)
$$\begin{aligned} D_1 f(c) &= \lambda_1 D_1 g_1(c) + \cdots + \lambda_k D_1 g_k(c), \\ &\cdots\cdots\cdots\cdots \\ D_p f(c) &= \lambda_1 D_p g_1(c) + \cdots + \lambda_k D_p g_k(c). \end{aligned}$$

DEMOSTRACION. Si se aplica la fórmula (42.7) a $e_1, \ldots, e_p \in \boldsymbol{R}^p$, se obtiene un sistema de ecuaciones con el lado derecho de (42.9) y el lado izquierdo de (42.9) multiplicados por μ. Si $\mu = 0$,entonces el supuesto de que el rango es igual a K implica que $\lambda_1 = \cdots = \lambda_k = 0$, contrario a la hipótesis. Por lo tanto, $\mu \neq 0$ y se puede normalizar este sistema para obtener (42.9).
Q.E.D.

42.11 EJEMPLO. Encontrar los puntos en la intersección del cilindro $\{(x, y, z): x^2 + y^2 = 4\}$ y el plano $\{(x, y, z): 6x + 3y + 2z = 6\}$ que sean los más cercanos al origen y aquellos que sean los más lejanos al origen

Se buscarán extremos relativos de la función

$$f(x, y, z) = x^2 + y^2 + z^2$$

sujeta a las restricciones

$$g_1(x, y, z) = x^2 + y^2 - 4 = 0,$$
$$g_2(x, y, z) = 6x + 3y + 2z - 6 = 0.$$

La matriz que corresponde a (42.8) en este caso es

$$\begin{bmatrix} 2x & 6 \\ 2y & 3 \\ 0 & 2 \end{bmatrix}$$

que tiene rango 2 excepto en el punto $(x, y) = (0, 0)$ que no satisface las restricciones. Por lo tanto, puede aplicar el corolario para obtener el sistema.

$$\begin{aligned} 2x &= \lambda_1(2x) + \lambda_2(6), \\ 2y &= \lambda_1(2y) + \lambda_2(3), \\ 2z &= \qquad\quad \lambda_2(2), \\ x^2 + y^2 &= 4, \\ 6x + 3y + 2z &= 6, \end{aligned}$$

de cinco ecuaciones en cinco variables. La tercera ecuación dá $\lambda_2 = z$, de tal manera que se puede substituir λ_2 de las dos primeras ecuaciones. Para substituir a λ_1, se multiplica la primera ecuación resultante por y y la segunda por x y se resta para obtener

$$0 = 6yz - 3xz = 3z(2y - x).$$

Se infiere que $z = 0$ o bien $x = 2y$.

Si $z = 0$, la quinta ecuación da $2x + y = 2$. Al combinarlo con la cuarta ecuación se obtiene

$$x^2 + (2 - 2x)^2 = x^2 + 4 - 8x + 4x^2 = 4,$$

por lo que $5x^2 - 8x = x(5x - 8) = 0$ y entonces $x = 0$ ó $x = 8/5$. Este caso nos lleva a los

dos puntos (0,2,0) y $(8/5, -6/5, 0)$ cada uno de los cuales tiene distancia 2 desde el origen.

Por otro lado, si $x = 2y$, la cuarta ecuación da $5y^2 = 4$ de tal manera que $y = 2/\sqrt{5}$ (y $x = 4/\sqrt{5}$) o $y = -2/\sqrt{5}$ (y $x = -4/\sqrt{5}$). Substituyendo en la quinta ecuación se obtiene $z = 3(1-\sqrt{5})$ y $z = 3(1+\sqrt{5})$, respectivamente. Por lo tanto, este caso nos lleva a los dos puntos $(4/\sqrt{5}, 2/\sqrt{5}, 3(1-\sqrt{5}))$ y $(-4/\sqrt{5}, -2/\sqrt{5}, 3(1+\sqrt{5}))$. Los cuadrados de las distancias entre estos puntos y el origen se puede ver que son $58 - 18\sqrt{5}$ y $58 + 18\sqrt{5}$, respectivamente.

Se deduce que ambos puntos (0,2,0) y $(8/5, -6/5, 0)$ minimizan la distancia desde el origen a esta intersección y que el punto $(-4/\sqrt{5}, -2/\sqrt{5}, 3(1+\sqrt{5}))$ maximiza tal distancia. Por consideraciones geométricas también se puede ver que el punto $(4/\sqrt{5}, 2/\sqrt{5}, 3(1-\sqrt{5}))$ da un máximo relativo entre puntos de esta intersección. (El lector deberá hacer un diagrama para poder visualizar mejor esta situación.)

Restricciones de desigualdad

En los últimos años los problemas sobre extremo que comprenden restricciones que son desigualdades más que igualdades han venido siendo cada vez más importantes. De modo que puede ser deseable encontrar un extremo relativo de una función $f: \Omega \to \boldsymbol{R}$ entre todos los puntos en $\Omega \subseteq \boldsymbol{R}^p$ que satisfaga las restricciones

$$h_1(x) \geq 0, \ldots, h_k(x) \geq 0.$$

Se podrá ver que dichos problemas también se pueden manejar por medio del método del Lagrange.

En algunas ocasiones un problema sobre extremo puede comprender tanto igualdades como desigualdades, pero puesto que la igualdad $g(x) = 0$ es equivalente a la desigualdad $-(g(x))^2 \geq 0$, dichos problemas siempre se podrán reducir a uno que comprenda solamente restricciones de desigualdad.

42.12 TEOREMA. *Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y suponga que f y $h_1, \ldots, h_k$ son funciones de valor real en $C^1(\Omega)$. Suponga que $c \in \Omega$ satisface las restricciones de desigualdad.*

$$h_1(x) \geq 0, \ldots, h_k(x) \geq 0,$$

y que existe una vecindad abierta U de c tal que $f(x) \leq f(c)$ [o $f(x) \geq f(c)$] para toda $x \in U$ que satisfaga estas restricciones. Entonces, existen números reales $\mu, \lambda_1, \ldots, \lambda_k$ no todos igual a cero tales que

$$(42.10) \qquad \mu Df(c) = \lambda_1 Dh_1(c) + \cdots + \lambda_k Dh_k(c).$$

Más aún, si $h_i(c) > 0$ para alguna i, entonces

DEMOSTRACION. Defínase $F: U \to \boldsymbol{R}^{k+1}$ como

$$F(x) = (f(x), h_1(x), \ldots, h_k(x)) \qquad \text{para } x \in U.$$

Si c es un punto en U en donde las restricciones se satisfacen y en donde f es maximizado o bien minimizado, entonces $DF(c)$ no puede ser suprayectiva y entonces (42.10) se debe cumplir.

Si $h_1(c) = 0, \ldots, h_r(c) = 0$, pero $h_{r+1}(c) > 0, \ldots, h_k(c) > 0$, entonces sea $U_1 \subseteq U$ una vecindad abierta de c en la que $h_{r+1}, \ldots, h_k$ son estrictamente positivos y aplicar el teorema a las restricciones $h_1(x) \geq 0, \ldots, h_r(x) \geq 0$.

Q.E.D.

42.13 COROLARIO. *Además de las hipótesis del teorema 42.12 suponga que el rango de la matriz*

$$\begin{bmatrix} D_1 h_1(c) & \cdots & D_1 h_r(c) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ D_p h_1(c) & \cdots & D_p h_r(c) \end{bmatrix} \tag{42.11}$$

que corresponde a aquellos h_i *para los cuales* $h_i(c) = 0$, *es igual a r. Entonces se puede tomar* $\mu = 1$ *en (42.10). Además, si* $f(x) \leq f(c)$ [*respectivamente,* $f(x) \geq f(c)$] *para toda* $x \in U$ *que satisfaga las restricciones y si se toma* $\mu = 1$ *en (42.10), entonces* $\lambda_i \leq 0$ [*respectivamente,* $\lambda_i \geq 0$] *para* $i = 1, \ldots, r$.

DEMOSTRACION. La demostración de que se puede tomar $\mu = 1$ en (42.10) es análoga a la del corolario 42.10. Suponga entonces que $\mu = 1$ y que $f(x) \leq f(c)$ para toda $x \in U$ que satisfaga las restricciones. Dado que el rango de la matriz (42.11) es $r \leq k$, entonces si $1 \leq j \leq r$ existe un vector $v_j \in \boldsymbol{R}^p$ tal que

$$Dh_i(c)(v_j) = \delta_{ij}.$$

Por lo tanto, si $t > 0$ es suficientemente pequeña, entonces en el segmento que une a c y $c + tv_j$ existe un punto c_t tal que

$$0 \geq f(c + tv_j) - f(c) = Df(c_t)(tv_j) = tDf(c_t)(v_j).$$

En consecuencia, se tiene

$$0 \geq \lim_{\substack{t \to 0 \\ t > 0}} \frac{f(c + tv_j) - f(c)}{t} = Df(c)(v_j) = \sum_{i=1}^{r} \lambda_i Dh_i(c)(v_j) = \lambda_j.$$

Por lo tanto, $\lambda_j \leq 0$ para $j = 1, \ldots, r$. Q.E.D.

Para una demostración elemental pero muy distinta del teorema de Lagrange que comprende restricciones de desigualdad, véase el artículo de E. J. Mc.Shane† citado en la lista de referencias.

† E.J. McShane (1904-) obtuvo su grado de doctor en la Universidad de Chicago. Ha estado relacionado durante mucho tiempo con la Universidad de Virginia y se le conoce ampliamente por sus contribuciones a la teoría de integración, el cálculo de variaciones' la teoría del control óptimo y la balística exterior.

Ejercicios

42.A Encontrar los puntos críticos de las siguientes funciones y determinar la naturaleza de estos puntos.

(a) $f(x, y) = x^2 + 4xy$,
(b) $f(x, y) = x^4 + 2y^4 + 32x - y + 17$,
(c) $f(x, y) = x^2 + 4y^2 - 12y^2 - 36y$,
(d) $f(x, y) = x^4 - 4xy$,
(e) $f(x, y) = x^2 + 4xy + 2y^2 - 2y$,
(f) $f(x, y) = x^2 + 3y^4 - 4y^3 - 12y^2$.

42.B. Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto, suponga que $f: \Omega \to \boldsymbol{R}$ tiene segundas derivadas parciales continuas en Ω, que $c \in \Omega$ es un punto crítico de f y que $\delta > 0$.

(a) Demostrar que si $D^2 f(x)(w)^2 \geq 0$ para toda $0 < \|x - c\| < \delta$ y $w \in \boldsymbol{R}^p$, entonces c es un punto de mínimo relativo de f.

(b) Demostrar que si $D^2 f(x)(w)^2 > 0$ para toda $0 < \|x - c\| < \delta$ y $w \in \boldsymbol{R}^p$, $w \neq 0$, entonces c es un punto de mínimo relativo estricto de f.

42.C. Sea $\Omega \subseteq \boldsymbol{R}^2$ abierto, suponga que $f: \Omega \to \boldsymbol{R}$ tiene segundas derivadas parciales continuas en Ω, que $c \in \Omega$ es un pun punto crítico de f y que

$$\Delta(x) = D_{11} f(x) D_{22} f(x) - (D_{12} f(x))^2$$

para $x \in \Omega$. Suponga además que para alguna $\delta > 0$, $\Delta(x) \geq 0$ para toda $\|x - c\| < \delta$.

(a) Si $D_{11} f(x) > 0$ (o si $D_{22} f(x) > 0$) para toda x tal que $0 < \|x - c\| < \delta$, demostrar que c es un punto de mínimo relativo de f.

(b) Si $D_{11} f(x) < 0$ (o si $D_{22} f(x) < 0$) para toda x tal que $0 < \|x - c\| < \delta$, demostrar que c es un punto de máximo relativo de f.

42.D. Sea $f: \boldsymbol{R}^p \to \boldsymbol{R}$ diferenciable en $\boldsymbol{R}^p$ y $f(x) = 0$ para toda $x \in \boldsymbol{R}^p$ con $\|x\| = 1$. Demostrar que existe un punto $c \in \boldsymbol{R}^p$ con $\|c\| < 1$ tal que $Df(c) = 0$. (Esta es una versión del teorema de Rolle en $\boldsymbol{R}^p$.)

42.E. Usar el teorema de transformación suprayectiva 41.6 para demostrar el corolario 42.2.

42.F. Probar que cada una de las siguientes funciones tiene un punto crítico en el origen. Encontrar cuáles tienen extremos relativos y cuáles tienen puntos silla en el origen.

(a) $f(x, y) = x^2 y^2$,
(b) $f(x, y) = x^2 - y^3$,
(c) $f(x, y) = x^3 - y^3$,
(d) $f(x, y) = x^4 - x^2 y^2 + y^4$,
(e) $f(x, y) = x^3 y - x y^3$,
(f) $f(x, y) = x^4 + y^4$.

42.G. Demostrar que la función $f(x, y) = 2x + 4y - x^2 y^4$ tiene un punto crítico pero no tiene puntos extremos relativos.

42.H. Estudiar el comportamiento de la función $f(x, y) = x^3 - 3xy^2$ (2, 1, −3) en una vecindad del origen. A la gráfica de esta función algunas veces se la llama "silla de mono". ¿Por qué?

42.I. Encontrar la distancia mínima del punto $2x + y -$ al plano $2z = 4$.

42.J. Encontrar las dimensiones de la caja rectangular, abierta por arriba, con volumen dado y superficie mínima.

42.K. Encontrar la distancia mínima entre las rectas $L_1 = \{(x, y, z): x = 2 - t,\ y = 3 + t,\ z = 1 - 2t\}$ y $L_2 = \{(x, y, z): x = 1 - s,\ y = 2 - s,\ z = 3 + s\}$.

42.L. Dar ejemplos para probar que es falsa cada una de las conjeturas establecidas después del teorema 42.5

42.M. Suponga que se tienen n puntos (x_j, y_j), $j = 1, \ldots, n$ en $\boldsymbol{R}^2$ y se desea encontrar la función afín $F: \boldsymbol{R} \to \boldsymbol{R}$ dada por $F(x) = Ax + B$, tal que la cantidad

$$\sum_{j=1}^{n} (F(x_j) - y_j)^2$$

se minimice. Demostrar que esto da origen a las ecuaciones

$$A \sum_{j=1}^{n} x_j^2 + B \sum_{j=1}^{n} x_j = \sum_{j=1}^{n} x_j y_j,$$
$$A \sum_{j=1}^{n} x_j + nB = \sum_{j=1}^{n} y_j,$$

para los números A, B. [Esta función F se dice que es la función afín que "mejor le queda a los n puntos en el sentido de **mínimos cuadrados**".]

42.N. Sea $f: [0, 1] \to \boldsymbol{R}$ continua en $[0, 1]$. Se desea escoger números reales A, B, C de tal manera que minimicen la cantidad

$$\int_0^1 [f(x) - (Ax^2 + Bx + C)]^2 \, dx.$$

Demostrar que se pueden elegir A, B, C tales que satisfagan el sistema

$$\tfrac{1}{5}A + \tfrac{1}{4}B + \tfrac{1}{3}C = \int_0^1 x^2 f(x) \, dx,$$

$$\tfrac{1}{4}A + \tfrac{1}{3}B + \tfrac{1}{2}C = \int_0^1 x f(x) \, dx,$$

$$\tfrac{1}{3}A + \tfrac{1}{2}B + C = \int_0^1 f(x) \, dx.$$

[La función que resulta $x \mapsto Ax^2 + Bx + C$ se dice que es la función cuadrática que "mejor le queda a f en $[0, 1]$ en el sentido de **mínimos cuadrados**".]

42.O. Usar el teorema de Lagrange para localizar puntos en la curva $y = x^5 + x - 2$ en donde la función $f(x, y) = x - y$ pueda tener algún extremo relativo. Después, trazar la curva y las curvas de nivel de f para probar que los punto(s) que se localicen no son punto(s) de extremos relativos de f.

42.P. Sea $f: \boldsymbol{R}^2 \to \boldsymbol{R}$ la función cuadrática $f(x, y) = ax^2 + 2bxy + cy^2$ para $(x, y) \in \boldsymbol{R}^2$. Se desea encontrar los extremos relativos de f en el círculo unitario $\{(x, y): x^2 + y^2 = 1\}$. Usar el teorema de Lagrange para probar que los puntos (x_0, y_0) en donde estos extremos relativos se toman deben satisfacer el sistema

$$(a - \lambda)x_0 + by_0 = 0,$$
$$bx_0 + (c - \lambda)y_0 = 0,$$

en donde el multiplicador de Lagrange λ es una raíz de la ecuación

$$\lambda^2 - (a + c)\lambda + (ac - b^2) = 0.$$

Demostrar que el valor correspondiente del multiplicador λ es igual al valor extremo de f en dicho extremo relativo.

42.Q. La suma de tres números reales es 9. Encontrar estos números si su producto se ha de maximizar.

42.R. Demostrar que el volumen de la caja más grande que se puede inscribir en la región elipsoidal

$$\left\{(x, y, z): \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \leq 1\right\}$$

(en donde a,b,c son números estrictamente positivos) es igual a $8\,abc/3\sqrt{3}$.

42. S. Para cada una de las siguientes funciones encontrar los valores máximo y mínimo en el conjunto dado. (Cuando sea conveniente, tomar en cuenta los signos de los multiplicadores.)

(a) $f(x, y) = x^4 - y^4$, $\quad S = \{(x, y): x^2 + y^2 \leq 1\}$.
(b) $f(x, y) = x^2 + 2x + y^2$, $\quad S = \{(x, y): x^2 + y^2 \leq 1\}$.
(c) $f(x, y) = x^2 + 2x + y^2$, $\quad S = \{(x, y), |x| \leq 1, |y| \leq 1\}$.
(d) $f(x, y) = (1 - x^2) \sin y$, $\quad S = \{(x, y), |x| \leq 1, |y| \leq \pi\}$.

42.T. Defina f de $x > 0$, $y > 0$ a $\boldsymbol{R}$ como $f(x, y) = 1/x + cxy + 1/y$.

(a) Localizar los puntos críticos de f y determinar su naturaleza.

(b) Si $c > 0$, sea $S = \{(x, y): 0 < x,\ 0 < y,\ x + y \leq c\}$. Encontrar los valores máximo relativo y mínimo relativo de f en S.

42.U. Encontrar los valores extremos de $f(x, y, z) = x^3 + y^3 + z^3$ sujeta a las restricciones $x^2 + y^2 + z^2 = 1$ y $x + y + z = 1$.

42.V. Supóngase que f tiene segundas derivadas parciales continuas en un conjunto abierto que contiene a la bola $\{x \in \boldsymbol{R}^p : \|x\| \leq r\}$ a $\boldsymbol{R}$ y suponga que existe c con $\|c\| < r$ tal que

$$M = f(c) > \sup\{f(x): \|x\| = r\} = m.$$

Defina g como

$$g(x) = f(x) + \frac{M - m}{4r^2}\|x - c\|^2.$$

Demostrar que $g(c) = M$, mientras que $g(x) < M$ para $\|x\| = r$. Por lo tanto, g alcanza un máximo relativo en algún punto c_1 con $\|c_1\| < r$, en donde se tiene

$$\sum_{j=1}^{p} D_{jj} f(c_1) \leq -\frac{p}{2r^2}(M - m) < 0.$$

42.W. Sea $\Omega \subseteq \boldsymbol{R}^p$ un conjunto abierto acotado, sea $b(\Omega)$ el c conjunto de puntos frontera de Ω (véase la definición 9.7) y sea $\Omega^- = \Omega \cup b(\Omega)$ la cerradura de Ω. Se dice que una función $f: \Omega^- \to \boldsymbol{R}$ es **armónica** en Ω si es continua en Ω^- y satisface la ecuación de Laplace

$$\sum_{j=1}^{p} D_{jj} f(x) = 0$$

para toda $x \in \Omega$.

(a) Usar el argumento del ejercicio anterior para probar que una función armónica en Ω alcanza su supremo e ínfimo en $b(\Omega)$.

(b) Si f y g son armónicas en Ω y si $f(x) = g(x)$ para $x \in b(\Omega)$, entonces $f(x) = g(x)$ para toda $x \in \Omega^-$.

(c) Si f y g son armónicas en Ω $f(x) = \varphi(x)$, $g(x) = \psi(x)$ para $x \in b(\Omega)$, entonces

$$\sup\{|f(x)-g(x)|: x\in\Omega\} = \sup\{|\varphi(x)-\psi(x)|: x\in b(\Omega)\}.$$

(Esta conclusión se puede establecer diciendo que "las soluciones del problema de Dirichlet para Ω dependen continuamente de los datos de la frontera".)

42.X. Probar que el máximo de $f(x_1, \ldots, x_p) = (x_1 \cdots x_p)^2$ sujeta a la restricción $x_1^2 + \cdots + x_p^2 = 1$ es igual a $1/p^p$. Usar esto para obtener la desigualdad:

$$|y_1 \cdots y_p| \le \frac{\|y\|^p}{p^{p/2}} \quad \text{para } y \in \boldsymbol{R}^p.$$

42.Y. Demostrar que la media geométrica de la colección de números reales positivos $\{a_1, \ldots, a_p\}$ no excede a su medida aritmética, es decir

$$(a_1 \cdots a_p)^{1/p} \le \frac{1}{p}(a_1 + \cdots + a_p).$$

42.Z. *(a)* Sean $p>1$, $q>1$, $1/p+1/q=1$. Probar que el mínimo de $f(x, y) = (1/p)x^p + (1/q)y^q$ $(x>0, y>0)$ sujeta a la restricción $xy = 1$ es igual a 1.

(b) Usar la parte *(a)* para probar que si $a>0$, $b>0$, entonces

$$ab \le \frac{1}{p}a^p + \frac{1}{q}b^q.$$

(c) Sean $\{a_i\}$, $\{b_i\}$, $i = 1, \ldots, n$, números reales positivos. Demostrar la desigualdad de Holder:

$$\sum_{i=1}^n a_i b_i \le \left(\sum_{i=1}^n a_i^p\right)^{1/p}\left(\sum_{i=1}^n b_i^q\right)^{1/q}$$

tomando $A = (\sum a_i^p)^{1/p}$, $B = (\sum b_i^q)^{1/q}$, y aplicando la parte *(c)* $a = a_i/A$, $b = b_i/B$.

(d) Usar la desigualdad de Holder en *(c)* para obtener la desigualdad de Minkowski:

$$\left(\sum_{i=1}^n |a_i+b_i|^p\right)^{1/p} \le \left(\sum_{i=1}^n |a_i|^p\right)^{1/p} + \left(\sum_{i=1}^n |b_i|^{1/p}\right)^{1/p}$$

[*Sugerencia:* $|a+b|^p = |a+b|\,|a+b|^{p/q} \le |a|\,|a+b|^{p/q} + |b|\,|a+b|^{p/q}$.]

VIII INTEGRACION EN $\mathbf{R}^p$

En este capítulo se estudia la teoría de integración de funciones de valor real en $\mathbf{R}^p$ en donde $p > 1$. La vía de acceso que aquí se usa es la misma que se estableció en la sección 29 para el caso p = 1, pero aquí sólo se trata la integral de Riemann (y no la integral de Riemann-Stieltjes).

En la sección 43 se verá que para funciones acotadas definidas en una celda cerrada en $\mathbf{R}^p$, la teoría prácticamente no cambia de como era $\mathbf{R}$. Sin embargo, con el objeto de poder integrar sobre conjuntos más generales en $\mathbf{R}^p$ es necesario desarrollar teoría de "contenido" (como se llamará al concepto p-dimensional de "área") para una familia apropiada de conjuntos en $\mathbf{R}^p$. como se hace en la sección 44. Se habrá de caracterizar la función contenido en esta familia de conjuntos y se mostrará la manera de expresar integrales en $\mathbf{R}^p$ como integrales iteradas. La última sección está dedicada al desarrollo de teoremas importantes acerca de las transformaciones de conjuntos e integrales en aplicaciones diferenciales. Las dificultades teóricas son considerables, pero se concluye con un teorema muy útil que justifica el cambio de variables en casos en que la transformación pueda poseer una cantidad limitada de comportamieto "peculiar".

Sección 43 La integral en $\mathbf{R}^p$

En las secciones 29-31 se analizó la integral de funciones acotadas de valor real definidas en un intervalo compacto J en $\mathbf{R}$. El lector capaz de captar generalizaciones habrá observado que una parte considerable de lo que se hizo en aquellas secciones se cumple cuando los valores de la función están en un espacio cartesiano $\mathbf{R}^q$. Una vez que se reconoce esta posibilidad no es difícil llevar a cabo las modificaciones necesarias para obtener una teoría de integración para funciones de J a $\mathbf{R}^q$.

También es natural preguntar si se puede obtener una teoría de integración para funciones cuyo dominio sea un subconjunto del espacio $\mathbf{R}^p$, y el lector recordará que, de hecho, esto se hizo en cursos de cálculo en que se consideraban integrales "dobles" y "triples". En esta sección se habrá de iniciar un

estudio de la integral de Riemann de funciones de valor real definidas en un subconjunto acotado apropiado de $\mathbf{R}^p$. Aun cuando muchos de los resultados se pueden extender de tal manera que los valores estén en $\mathbf{R}^q$ para $q > 1$, esta extensión se le deja al lector.

Contenido cero

De la sección 5 recuerde que una celda en R es un conjunto que tiene u·· de las cuatro formas:

$$(a, b), \qquad [a, b], \qquad [a, b), \qquad (a, b],$$

en donde $a \leq b$. Los números a, b se llaman **puntos extremos** de estas celdas. Una **celda** en R^p es el producto cartesiano $J = J_1 \times \cdots \times J_p$ de p celdas en R. Se dice que la celda en **cerrada** (respectivamente, **abierta**) si cada una de las celdas $J_1, \ldots, J_p$ es cerrada (respectivamente, abierta) en R. Si las celdas J_i tienen puntos extremos $a_i \leq b_i (i = 1, \ldots, p)$, se define el **contenido** de $J = J_1 \times \cdots \times J_p$ como el producto

$$c(J) = (b_1 - a_1) \cdots (b_p - a_p).$$

Si $p = 1$, al contenido por lo general se le llama "longitud"; si $p = 2$, al contenido se le llama "área"; si $p = 3$, al contenido se le llama "volumen". Se usará la palabra "contenido" porque está libre de connotaciones especiales que estas otras palabras pueden tener.

Observe que si $J = J_1 \times \cdots \times J_p$, y $K = K_1 \times \cdots \times K_p$ son celdas en R^p tales que los puntos extremos de J_i y de K_i son los mismos para cada $i = 1, \ldots, p$, entonces $c(J) = c(K)$. En forma análoga, si $a_k = b_k$ para alguna $k = 1, \ldots, p$, entonces la celda J tiene contenido $c(J) = 0$; sin embargo, no es necesario que $J = \emptyset$. Si J_i es una celda con puntos extremos $a_i \leq b_i$ y si $b_1 - a_1 = \cdots = b_p - a_p > 0$, entonces se dice que

$$J = J_1 \times \cdots \times J_p$$

es un **cubo.** El cubo puede ser cerrado, abierto o ninguno de los dos. Al número $b_1 - a_1 > 0$ se le llama la **longitud lateral** del cubo.

43.1 DEFINICION. Un conjunto $Z \subseteq R^p$ tiene **contenido cero** si para cada $\varepsilon > 0$ existe un conjunto *finito* $J_1, \ldots, J_n$ de celdas cuya unión contiene a Z y tal que

$$c(J_1) + \cdots + c(J_n) < \varepsilon.$$

Es importante que el lector pruebe que aun cuando se puede pedir que las celdas que aparecen en esta definición sean cerradas, sean abiertas o sean cubos, el concepto de contenido cero sigue siendo exactamente el mismo.

43.2 EJEMPLOS. *(a)* Un punto en $\boldsymbol{R}^p$ tiene contenido cero. (¿por qué?) Más general, cualquier subconjunto finito de $\boldsymbol{R}^p$ tiene contenido cero.

(b) Si $(z_n)_{n\in N}$ es una sucesión en $\boldsymbol{R}^p$ que converge a $z_0 \in \boldsymbol{R}^p$, entonces el conjunto $Z = \{z_n : n \geq 0\}$ tiene contenido cero. Ya que si $\varepsilon > 0$ sea J_0 una celda abierta que contiene a z_0 tal que $c(J_0) < \varepsilon$. Por lo tanto, existe $k \in \boldsymbol{N}$ tal que $z_n \in J_0$ para toda $n > k$, y se puede tomar $J_i = \{z_i\}$ para $i = 1, \ldots, k$ para obtener $Z \subseteq J_0 \cup J_1 \cup \cdots \cup J_k$. Dado que

$$c(J_0) + c(J_1) + \cdots + c(J_k) < \varepsilon + 0 + \cdots + 0 = \varepsilon,$$

y como $\varepsilon > 0$ es arbitraria, se infiere que Z tiene contenido cero.

(c) Cualquier subconjunto de un conjunto con contenido cero tiene contenido cero. La unión de un número finito de conjuntos con contenido cero tiene contenido cero.

(d) En el espacio $\boldsymbol{R}^2$, el conjunto con forma de diamante $S = \{(x, y) : |x| + |y| = 1\}$ tiene contendio cero. Puesto que si $n \in \boldsymbol{N}$, se introducen cuadrados con diagonales a lo largo de S y vértices en los puntos $x = y = \pm k/n$ $(k = 0, 1, \ldots, n)$, se puede ver entonces que S se puede encerrar en $4n$ cuadrados cerrados cada uno con contenido $1/n^2$. Por lo tanto, el contenido total es $4/n$, que se puede hacer arbitrariamente pequeño. (Véase la figura 43.1).

(e) El círculo $S = \{(x, y) : x^2 + y^2 = 1\}$ en $\boldsymbol{R}^2$ tiene contenido cero Esto se puede probar modificando el argumento en *(d)*.

(f) Sea f una función continua de $J = [a, b]$ a $\boldsymbol{R}$. Entonces la gráfica

$$G = \{(x, f(x)) \in \boldsymbol{R}^2 : x \in J\}$$

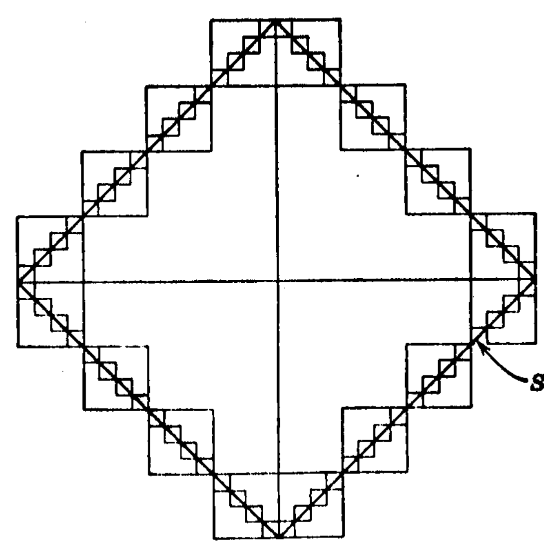


Figura 43.1

tiene contenido cero. Esto se puede probar usando la continuidad uniforme de f y modificando el argumento en (d).

(g) El conjunto $S \subseteq \boldsymbol{R}^2$ que consta de todos los puntos (x, y) en donde x y y pertencen a $\boldsymbol{I} \cap \boldsymbol{Q}$ es un conjunto contable pero no tiene contenido cero. De hecho, cualquier unión finita de celdas que contienen a S también debe contener a la celda $\boldsymbol{I} \times \boldsymbol{I}$, que tiene contenido 1.

A diferencia de *(f)*, se puede ver que *hay "curvas continuas" en* $\mathbf{R}^2$ *que tienen contenido positivo.* de hecho, hay funciones continuas f, g de $\boldsymbol{I} = [0, 1]$ a $\boldsymbol{R}$ tales que el conjunto.

$$S = \{(f(t), g(t)); t \in \boldsymbol{I}\}$$

contiene a la celda $\boldsymbol{I} \times \boldsymbol{I}$ en $\boldsymbol{R}^2$. A dicha curva se le llama **curva de espacio de rellenamiento** o **curva de Peano.** (Véase el ejercicio 43.U.)

Definición de la integral

Se definirá primero la integral para una función acotada f definida en una *celda cerrada* $I \subseteq \mathbf{R}^p$ y con valores en R. Sea

$$I = [a_1, b_1] \times \cdots \times [a_p, b_p],$$

y para cada $k = 1, \ldots, p$, sea P_k una partición de $[a_k, b_k]$ en un número finito de celdas cerradas en $\boldsymbol{R}$. Esto induce una **partición** P de I en un número finito de celdas cerradas en $\boldsymbol{R}^p$. Si P y Q son particiones de I, se dice que P es un **refinamiento** de Q si cada celda en P está contenida en alguna celda en Q. (De manera alternativa viendo que una partición está determinada por los vértices de sus celdas, P es un refinamiento de Q si y sólo si todos los vértices contenidos en Q también están en P.)

43.3 DEFINICION. Una **suma de Riemann** $S(P;f)$ correspondiente a una partición $P = \{J_1, \ldots, J_n\}$ de I está dada por

$$S(P; f) = \sum_{k=1}^{n} f(x_k)c(J_k),$$

en donde x_k es cualquier punto "intermedio" en J_k, $k = 1, \ldots, n$. Un número real L se define como la **integral de Riemann de** f **sobre** I si para toda $\varepsilon > 0$ hay una partición P_ε de I tal que si P es cualquier refinamiento de P_ε y $S(P;f)$ es cualquier suma de Riemann correspondiente a P, entonces $|S(P; f) - L| \le \varepsilon$. En caso de que esta integral exista se dice que f es **integrable sobre** I.

Es fácil demostrar que el valor L de la integral está determinado de manera única cuando existe; por lo general se escribirá

$$L = \int_I f;$$

sin embargo, cuando $p = 2$ en ocasiones la integral se denota por medio de

$$\iint_I f \qquad \text{ó} \qquad \iint_I f(x, y)\, d(x, y),$$

y cuando $p = 3$ en ocasiones se escribe

$$\iiint_I f \qquad \text{ó} \qquad \iiint_I f(x, y, z)\, d(x, y, z).$$

Hay un criterio de Cauchy para integrabilidad que resulta conveniente. Dado que su demostración es muy semejante a la del teorema 29.4, se omitirá.

43.4 CRITERIO DE CAUCHY. *Una función acotada* $f: I \to \mathbf{R}$ *es integrable en* I *si y sólo si para toda* $\varepsilon > 0$ *existe una partición* Q_ε *de* I *tal que si* P *y* Q *son particiones de* I *que son refinamientos de* Q_ε *y si* $S(P;f)$ *y* $S(Q;f)$ *son cualesquiera sumas de Reimann correspondientes; entonces,*

$$|S(P; f) - S(Q; f)| \leq \varepsilon.$$

Se desea considerar ahora funciones definidas en subconjuntos acotados de $\mathbf{R}^p$ que sean más generales que celdas cerradas. Sea $A \subseteq \mathbf{R}^p$ un conjunto *acotado* y sea $f: A \to \mathbf{R}$ una función *acotada*. Dado que A es acotado, existe una celda cerrada $I \subseteq \mathbf{R}^p$ tal que $A \subseteq I$. Se define $f_I: I \to \mathbf{R}$ como

$$\begin{aligned} f_I(x) &= f(x) \quad &&\text{para } x \in A, \\ &= 0 \quad &&\text{para } x \in I \setminus A. \end{aligned}$$

Si la función f_I es integrable en I en el sentido de la definición 43.3, entonces queda como ejercicio (véase 43.M) demostrar que el valor $\int_I f_I$ no depende de la elección de la celda cerrada I que contiene a A. Debido a esto, se dirá que f es **integrable en** A y se definirá

$$\int_A f = \int_I f_I,$$

ya que el lado derecho sólo depende de f y A. (En argumentos subsiguientes con frecuencia f_I se designará simplemente como f.)

Análogamente, sean A y B subconjuntos acotados de $\mathbf{R}^p$ y sea $f: A \to \mathbf{R}$. Sea I una celda cerrada que contiene a $A \cup B$ y defínase $f_1: I \to \mathbf{R}$ como

$$\begin{aligned} f_1(x) &= f(x) \quad &&\text{para } x \in A \cap B, \\ &= 0 \quad &&\text{para } x \in I \setminus (A \cap B). \end{aligned}$$

Observe que f_1 es la extensión a I de la restricción $f \mid A \cap B$. Si f_1 es integrable sobre I, se dice que f es **integrable en** B y se define

$$\int_B f = \int_I f_1 \quad \left(= \int_{A \cap B} f\right).$$

Propiedades de la integral

En seguida se darán algunas de las propiedades esperadas de la integral. A será un subconjunto acotado de $\mathbf{R}^p$.

43.4 TEOREMA. *Sean f y g funciones de A a R integrables en A y sean $\alpha, \beta \in \mathbf{R}$. Entonces, la función $\alpha f + \beta g$ es integrable en A y*

$$\int_A (\alpha f + \beta g) = \alpha \int_A f + \beta \int_A g.$$

DEMOSTRACION. Este resultado se infiere del hecho de que las sumas de Riemann para una partición P de una celda $I \supseteq A$ satisfacen

$$S(P; \alpha f + \beta g) = \alpha S(P; f) + \beta S(P; g),$$

cuando se usan los mismos puntos intermedios x_k. Q.E.D.

43.6 TEOREMA. *Si $f: A \to \mathbf{R}$ es integrable en A y si $f(x) \geq 0$ para $x \in A$, entonces $\int_A f \geq 0$.*

DEMOSTRACION. Observe que $S(P; f) \geq 0$ para cualquier suma de Riemann.

43.7 TEOREMA. *Sea $f: A \to \mathbf{R}$ una función acotada y suponga que A tiene contenido cero. Entonces, f es integrable en A y $\int_A f = 0$.*

DEMOSTRACION. Sea I una celda cerrada que contiene a A. Si $\varepsilon > 0$ está dada, sea P_ε una partición de I tal que aquellas celdas en P_ε que contengan puntos de A tengan contenido total menor que ε. (Demostrar que existe dicha partición P_ε.) Ahora, si P es un refinamiento de P_ε, entonces aquellas celdas en P que contienen puntos de A también tendrán contenido total menor que ε. Por lo que, si $|f(x)| \leq M$ para $x \in A$, se tiene $|S(P; f)| \leq M\varepsilon$ para cualquier suma de Riemann correspondiente a P. Dado que $\varepsilon > 0$ es arbitraria, esto implica que $\int_A f = 0$. Q.E.D.

43.8 TEOREMA. *Sean $f, g: A \to R$ funciones acotadas y supóngase que f es integrable en A. Supóngase que $E \subseteq A$ tiene contenido cero y que $f(x) = g(x)$ para toda $x \in A \setminus E$. Entonces, g es integrable en A y*

$$\int_A f = \int_A g.$$

DEMOSTRACION. Extiéndanse f y g a funciones f_I, g_I definidas en una celda cerrada I que contiene a A. Las hipótesis implican que $h_I = f_I - g_I$ es acotada y es igual a cero excepto en E. Por el teorema 43.7 se deduce que h_I es integrable en I y el valor de su integral es 0. Aplicando el teorema 43.5 se infiere que $g_I = f_I - h_I$ es integrable en I y

$$\int_A g = \int_I g_I = \int_I (f_I - h_I) = \int_I f_I = \int_A f. \qquad \text{Q.E.D.}$$

Existencia de la integral

Es de esperar que si f es continua de una celda cerrada I a $\mathbf{R}$, entonces f es integrable en I. Se habrá de probar un resultado más fuerte que permite que f sea discontinua en el complemento de un conjunto con contenido cero.

43.9 TEOREMA DE INTEGRABILIDAD. *Sea $I \subseteq \mathbf{R}^p$ una celda cerrada y supóngase que $f: I \to \mathbf{R}$ es acotada. Si existe un subconjunto $E \subseteq I$ con contenido cero tal que f es continua en $I \setminus E$, entonces f es integrable en I.*

DEMOSTRACION. Sea $|f(x)| \leq M$ para toda $x \in I$ y sea $\varepsilon > 0$ dada. Entonces existe (¿por qué?) una partición P_ε de I tal que *(i)* las celdas en P_ε que contienen algunos puntos de E los contienen en su interior y *(ii)* estas celdas tienen contenido total menor que ε. La unión C de celdas cerradas en P_ε que no contienen puntos de E es un subconjunto compacto en el que f es continua. De acuerdo con el teorema de continuidad uniforme 23.3, la restricción de f es uniformemente continua en C. Reemplazando P_ε por un refinamiento, si fuera necesario, se puede suponer que si J_k es celda en P_ε que está contenida en C y si $x, y \in J_k$, entonces $|f(x) - f(y)| < \varepsilon$.

Supóngase ahora que P y Q son refinamientos de P_ε. Si $S'(P; f)$ y $S'(Q; f)$ designan las porciones de las sumas de Riemann extendidas sobre las celdas en C, entonces un argumento análogo al que se usó en la segunda parte de la demostración del teorema 30.1 da

$$\begin{aligned} |S'(P; f) - S'(Q; f)| &\leq |S'(P; f) - S'(P_\varepsilon; f)| + |S'(P_\varepsilon; f) - S'(Q; f)| \\ &\leq 2\varepsilon c(I). \end{aligned}$$

Análogamente, si $S''(P; f)$ y $S''(Q; f)$ designan la porción restante de las sumas de Riemann, entonces

$$|S''(P; f) - S''(Q; f)| \leq |S''(P; f)| + |S''(Q; f)| \leq 2M\varepsilon.$$

Por lo tanto se infiere que

$$|S(P; f) - S(Q; f)| \leq \varepsilon(2c(I) + 2M);$$

dado que $\varepsilon > 0$ es arbitraria, por el criterio de Cauchy se infiere que f es integrable en I.

Las condiciones necesarias y suficientes para la integrabilidad se darán en el ejercicio 43.P y el proyecto 44.α.

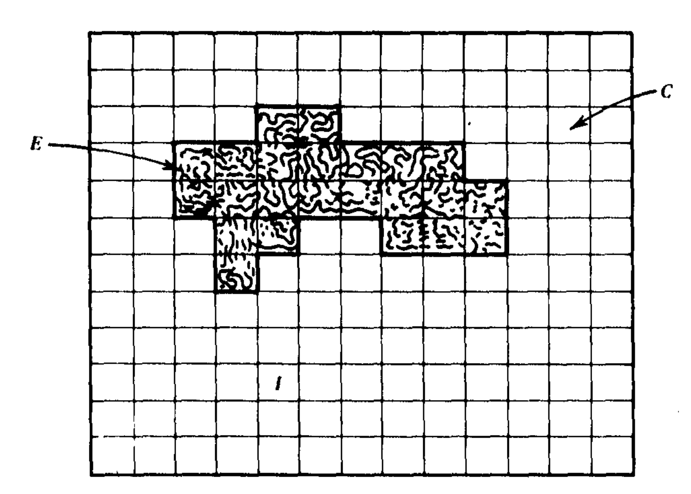


Figure 43.2

Ejercicios

43.A. *(a)* Sea $a = (a_1, \ldots, a_p) \in \boldsymbol{R}^p$ y sean $J_1 = [a_1, a_1], \ldots, J_p = [a_p, a_p]$ celdas en $\boldsymbol{R}$. Demostrar que $J = J_1 \times \cdots \times J_p$ tiene contenido cero en $\boldsymbol{R}^p$. De donde el conjunto $\{a\}$ tiene contenido cero en $\boldsymbol{R}^p$.

(b) Si se toma $J_1' = (a_1, a_1)$, entonces la celda $J' = J_1' \times J_2 \times \cdots \times J_p$ es vacía y tiene contenido cero.

43.B. Demostrar que un conjunto $Z \subseteq \boldsymbol{R}^p$ tiene contenido cero si y sólo si para cada $\varepsilon > 0$ existe un conjunto finito $K_1, \ldots, K_n$ de cubos cuya unión contiene a Z y tal que $c(K_1) + \cdots + c(K_n) < \varepsilon$.

43.C. Escribir los detalles de la demostración de la aseveración que se hizo en el ejemplo 43.2(*f*) acerca de que la gráfica $S \subseteq \boldsymbol{R}^2$ de una función continua $f : [a, b] \to \boldsymbol{R}$ tiene contenido cero.

43.D. Si J es una celda cerrada en $\boldsymbol{R}^2$ y $g : J \to \boldsymbol{R}$ es continua, demostrar que la gráfica $\{(x, y, g(x, y)) : (x, y) \in J)\} \subseteq \boldsymbol{R}^3$ de g tiene contenido a cero.

43.E. Sea $A \subseteq \boldsymbol{R}^2$ el conjunto que consta de todos los pares $(i/p, j/p)$ en donde p es un número primo y en donde $i, j = 1, 2, \ldots, p-1$. Demostrar que cada línea horizontal y cada línea vertical en $\boldsymbol{R}^2$ interseca a A en un número finito (a menudo cero) de puntos. ¿Tiene A contenido cero?

43.F. Sea $I \subseteq \boldsymbol{R}^p$ una celda cerrada y sean $P = \{I_1, \ldots, I_n\}$ y $Q = \{J_1, \ldots, J_m\}$ dos particiones de I en celdas cerradas. Demostrar que $R = \{I_i \cap J_j : i = 1, \ldots, n;\ j = 1, \ldots, m\}$ es una partición de I y que R es un refina-

miento tanto de P como de Q. A la partición R se le llama el **refinamiento común** de P y Q.

43.G. Si $I \subseteq J$ son celdas en $\boldsymbol{R}^p$ y si P es una partición de I, demostrar que existe una partición Q de J tal que toda celda en P pertenece a Q.

43.H. Sea $Z \subseteq \boldsymbol{R}^p$ un conjunto con contenido cero y sea I una celda cerrada que contiene a Z. Si $J_1, \ldots, J_n$ son celdas contenidas en I cuya unión contiene a Z, demostrar que existe una partición P de I tal que la cerradura de cada J_k es la unión de celdas en P.

43.I. Sea $Z \subseteq \boldsymbol{R}^p$ conjunto con contenido cero y sea I una celda cerrada que contiene a Z. Si $\varepsilon > 0$, demostrar que existe una partición P_ε de I tal que las celdas en P_ε que contienen puntos en Z tienen contenido total menor que ε.

43.J. En el ejercicio anterior, demostrar que se puede elegir P_ε con la propiedad adicional de que las celdas en P_ε que contengan puntos de Z los contengan en su interior.

43.K. En el ejercicio 43.I, si I es un cubo, demostrar que existe una partición Q de I en cubos tal que los cubos en Q que contienen puntos en Z tengan contenido total menor que ε.

43.L. Sea $A \subseteq \boldsymbol{R}^p$ acotado y sean I, J celdas cerradas en $\boldsymbol{R}^p$ tales que $A \subseteq I \subseteq J$. Si $f: A \to \boldsymbol{R}$ es una función acotada, defínase $f_I: I \to \boldsymbol{R}$ (respectivamente, $f_J: J \to \boldsymbol{R}$) no la función que coincide con f en A y es cero en I/A (respectivamente, J/A). Demostrar que la integral de f_I sobre I existe si y sólo si la integral de f_J sobre J existe en cuyo caso estas integrales son iguales.

43.M. Sea $A \subseteq \boldsymbol{R}^p$ acotado y sean I_1 e I_2 celdas cerradas en $\boldsymbol{R}^p$ tales que $A \subseteq I_i$. Sea $f: A \to \boldsymbol{R}$ una función acotada y para $i = 1, 2$, defínase $f_i: I_i \to \boldsymbol{R}$ como $f_i(x) = f(x)$ para $x \in A$ y $f_i(x) = 0$ para $x \in I_i \setminus A$. Demostrar que la integral del f_1 sobre I_1 existe si y sólo si la integral de f_2 sobre I_2 existe, en cuyo caso estas integrales son iguales.

43.N. Probar la unicidad de la integral de una función acotada f definida en una celda cerrada $I \subseteq \boldsymbol{R}^p$.

43.O. Escribir los detalles de la demostración del criterio de Cauchy 43.4.

43.P. Sea $I \subseteq \boldsymbol{R}^p$ una celda cerrada y sea $f: I \to \boldsymbol{R}$ acotada. Entonces f es integrable en I si y sólo si para toda $\varepsilon > 0$ existe una partición P_ε de I tal que si $P = \{J_1, \ldots, J_n\}$ es finameinto de P_ε, entonces

$$\sum_{j=1}^{n} (M_j - m_j) c(J_j) < \varepsilon,$$

en donde $M_j = \sup \{f(x): x \in J_j\}$ y $m_j = \inf \{f(x): x \in J_j\}$ para $j = 1, \ldots, n$. A este resultado se le llama el **criterio de Riemann para integrabilidad** (cf. teorema 30.1)

43.Q. Sea $I \subseteq \boldsymbol{R}^p$ una celda cerrada y sea $f: I \to \boldsymbol{R}$ acotada por M. Si f es integrable en I, demostrar que la función $|f|$ es integrable en I y que $\int_I |f| \leq Mc(I)$.

43.R. Sea $I \subseteq \boldsymbol{R}^p$ una celda cerrada y supóngase que $f, g: I \to \boldsymbol{R}$ son integrables en I. Demostrar que la función producto fg es integrable en I.

43.S. Sea $I \subseteq \boldsymbol{R}^p$ una celda cerrada y sea (f_n) una sucesión de funciones que son integrables en I. Si la sucesión converge uniformemente en I a f, demostrar que f es integrable en I y que

$$\int_I f = \lim_n \left(\int_I f_n \right).$$

43.T. Sea $K \subseteq \mathbf{R}^p$ un cubo cerrado y sean f, $g: K \to \mathbf{R}$ continuas. Demostrar que si $\varepsilon > 0$, entonces existe una partición $P_\varepsilon = \{K_1, \ldots, K_r\}$ de K en cubos tal que si x_j, y_j son cualesquier puntos de K_j, $j = 1, \ldots, r$, entonces

$$\left| \int_K fg - \sum_{j=1}^{r} f(x_j) g(y_j) c(K_j) \right| < \varepsilon c(K).$$

43.U. (Este ejercicio da un ejemplo que se debe a I. J. Schoenberg[1] de una curva de espacio de rellenamiento.) Sea $\varphi: \mathbf{R} \to \mathbf{R}$ continua, par, con periodo 2 y tal que

$$\begin{aligned} \varphi(t) &= 0 \quad &\text{para} \quad &0 \le t \le \tfrac{1}{3}, \\ &= 3t - 1 \quad &\text{para} \quad &\tfrac{1}{3} < t < \tfrac{2}{3}, \\ &= 1 \quad &\text{para} \quad &\tfrac{2}{3} \le t \le 1. \end{aligned}$$

(a) Dibujar la gráfica de φ. Obsérvese que $\|\varphi\|_{\mathbf{R}} = 1$.
(b) Si $t \in \mathbf{I}$, defínanse $f(t)$ y $g(t)$ como

$$f(t) = \frac{1}{2}\varphi(t) + \frac{1}{2^2}\varphi(3^2 t) + \frac{1}{2^3}\varphi(3^4 t) + \cdots,$$

$$g(t) = \frac{1}{2}\varphi(3t) + \frac{1}{2^2}\varphi(3^3 t) + \frac{1}{2^3}\varphi(3^5 t) + \cdots$$

Demostrar que estas series son uniformemente convergentes, de tal manera que f y g son continuas en $\mathbf{I}$.

(c) Calcular en donde $f(t)$, y $g(t)$, en donde t tiene la expansión ternaria (base 3) dada por

$$0.2020, \qquad 0.0220, \qquad 0.0022, \qquad 0.2002.$$

(d) Supóngase que (x, y) pertenece a la gráfica $S = \{(f(t), g(t)) : t \in \mathbf{I}\}$ y escríbanse x y y en sus expansiones binarias (base 2):

$$x = 0.\alpha_1\alpha_2\alpha_3\ldots, \qquad y = 0.\beta_1\beta_2\beta_3\ldots,$$

en donde α_n, β_m son 0 o bien 1. Sea t el número real correspondiente cuya expansión naria es

$$t = 0.(2\alpha_1)(2\beta_1)(2\alpha_2)(2\beta_2)\ldots.$$

Demostrar que $x = f(t)$ y $y = g(t)$. Por lo que todo punto en el cubo $\mathbf{I} \times \mathbf{I}$ pertenece a la gráfica S.

43.V. Un conjunto $Z \subseteq \mathbf{R}^p$ tiene **medida cero** si para cada $\varepsilon > 0$ hay una sucesión (J_n) de celdas cuya unión contiene a Z y tal que $\sum c(J_n) < \varepsilon$.

(a) Dado que el conjunto vacío es una celda, demostrar que un conjunto que tiene contenido cero también tiene medida cero.

(b) Demostrar que todo conjunto contable en $\mathbf{R}^p$ tiene medida cero. Por lo tanto, el conjunto en el ejemplo 43.2(g) tiene medida cero (pero no tiene contenido cero).

(c) Demostrar que en la definición "medida cero" que se acaba de dar se puede pedir que las celdas sean abiertas o que sean cubos.

[1]ISAAC J. SCHOENBERG (1903-) nació en Rumania y recibió su educación allí y en Alemania. Durante muchos años ha trabajado en la Universidad de Pensilvania en teoría de números, análisis real y complejo y en cálculo de variaciones.

(d) Demostrar que todo conjunto compacto con medida cero también tiene contenido cero.

(e) La unión de una familia contable de conjuntos con contenido cero tiene medida cero.

Proyectos

43.α. Sea $I \subseteq \mathbf{R}^p$ una celda cerrada y supóngase que $f: I \to \mathbf{R}$ es acotada. Si $P = \{J_1, \ldots, J_n\}$ es una partición de I, sea

$$m_j = \inf\{f(x) : x \in J_j\}, \qquad M_j = \sup\{f(x) : x \in J_j\}$$

para $j = 1, \ldots, n$ y defínanse las **sumas inferior** y **superior** de f para P como

$$L(P; f) = \sum_{j=1}^{n} m_j c(J_j), \qquad U(P; f) = \sum_{j=1}^{n} M_j c(J_j).$$

(a) Si $S(P; f)$ es cualquier suma de Riemann correspondiente a P, entonces $L(P; f) \leq S(P; f) \leq U(P; f)$. Si $\varepsilon > 0$, existen sumas de Riemann $S_1(P; f)$ y $S_2(P; f)$ correspondientes a P tales que

$$S_1(P; f) \leq L(P; f) + \varepsilon, \qquad U(P; f) - \varepsilon \leq S_2(P; f).$$

(b) Si P es una partición de I y Q es un refinamiento de P, entonces

$$L(P; f) \leq L(Q; f) \leq U(Q; f) \leq U(P; f).$$

(c) Si P_1 y P_2 son cualesquiera particiones de I, entonces $L(P_1; f) \leq U(P_2; f)$.

(d) Defínanse la **integral inferior** y **superior** de f en I como

$$L(f) = \sup\{L(P; f)\}, \qquad U(f) = \inf\{U(P; f)\},$$

respectivamente, en donde el supremo y el ínfimo se toman sobre todas las particiones de I. Demostrar que $L(f) \leq U(f)$.

(e) Demostrar que f es integrable (en el sentido de la definición 43.3) si y sólo si $L(f) = U(f)$, en cuyo caso $L(f) = U(f) = \int_I f$.

(f) Demostrar que f es integrable si y sólo si para cada $\varepsilon > 0$ existe una partición P tal que $U(P; f) - L(P; f) < \varepsilon$. A esta condición algunas veces se la llama condición de Riemann; compararla con el ejercicio 43.P.

43.B. Este proyecto desarrolla la integral de funciones en una celda cerrada $I \subseteq \mathbf{R}^p$ y con valores en $\mathbf{R}^q$. Si $P = \{J_1, \ldots, J_n\}$ es una partición de I, entonces una suma de Riemann $S(P; f)$ correspondiente a P es una suma

$$S(P; f) = \sum_{k=1}^{n} c(J_k) f(x_k)$$

en donde $x_k \in J_k$. Un elemento $L \in \mathbf{R}^q$ es la integral de Riemann de f sobre I si para toda $\varepsilon > 0$ existe una partición P_ε de I tal que si P es un refinamiento de P_ε y $S(P; f)$ es cualquier suma de Riemann correspondiente, entonces $\|S(P; f) - L\| < \varepsilon$. Examinar cuáles teoremas de esta sección siguen siendo válidos para funciones con valores en

i458 (1595x2495x2 pcx)

$\boldsymbol{R}^q$. Demostrar que $f: \boldsymbol{I} \to \boldsymbol{R}^q$ es integrable si y sólo si cada función $f_j = e_j \cdot f$, $j = 1, \ldots, q$, es integrable. (Aquí, $e_1, \ldots, e_q$ son los vectores de la base usual en $\boldsymbol{R}^q$.)

Sección 44 Contenido y la integral

En esta sección se introducirá la colección de conjuntos con contenido y se caracterizará la función contenido como una función de valor real definida en esta colección de conjuntos. Después, se obtendrán más propiedades de la integral sobre conjuntos con contenido y se demostrará cómo la integral se puede calcular como una "integral iterada".

44.1 DEFINICION. Si $A \subseteq \boldsymbol{R}^p$, recuérdese que un punto $x \in \boldsymbol{R}^p$ se dice que es un **punto frontera** de A si toda vecindad de x contiene tanto puntos de A como puntos de su complemento $\mathscr{C}(A)$. La **frontera** de A es el subconjunto de $\boldsymbol{R}^p$ que consta de todos los puntos frontera de A; se designará con $b(A)$.

Si $A \subseteq \boldsymbol{R}^p$, recuérdese que un punto de $\boldsymbol{R}^p$ es precisamente uno de los siguientes: es un punto interior de A, es un punto frontera de A o es un punto exterior de A. El **interior** A^0 consta de todos los puntos interiores de A, es un conjunto abierto en $\boldsymbol{R}^p$. Como se vio anteriormente, la **frontera** $b(A)$ consta de todos los puntos frontera de A, es un conjunto cerrado en $\boldsymbol{R}^p$. La **cerradura** A^- es la unión $A \cup b(A)$; es un conjunto cerrado en $\boldsymbol{R}^p$.

Por lo general se espera que la frontera de un conjunto sea pequeña; esto se debe a que estamos acostumbrados a pensar en rectángulos, círculos y otras figuras elementales. En el ejemplo 43.2(g) se muestra que la frontera de un conjunto contable en $\boldsymbol{R}^2$ puede ser igual a $\boldsymbol{I} \times \boldsymbol{I}$.

Conjuntos con contenido

Se definirá ahora el contenido de un subconjunto de $\boldsymbol{R}^p$ cuya frontera tiene contenido cero.

44.2 DEFINICION. Un conjunto acotado $A \subseteq \boldsymbol{R}^p$ cuya frontera $b(A)$ tiene contenido cero se dice que tiene contenido. La colección de todos los subconjuntos de $\boldsymbol{R}^p$ que tienen contenido se designará como $\mathscr{D}(\boldsymbol{R}^p)$. Si $A \in \mathscr{D}(\boldsymbol{R}^p)$ y si I es una celda cerrada que contiene a A, entonces la función g_I definida como

$$g_I(x) = 1 \quad \text{para } x \in A,$$
$$= 0 \quad \text{para } x \in I \setminus A,$$

es continua en I excepto posiblemente en puntos de $b(A)$. Por lo tanto, g es integrable en I y se define el **contenido** $c(A)$ de A como $\int_I g_I$. De modo que

$$c(A) = \int_I g_I = \int_A 1.$$

Obsérvese que si $J \subseteq \boldsymbol{R}^p$ es una celda, entonces su frontera consta de la unión de un número finito de "caras" que son celdas con contenido cero cada una. (Por ejemplo, si $J = [a, b] \times [c, d]$,) entonces $b(J)$ es la unión de las cuatro celdas

$$[a, b] \times [c, c], \qquad [a, b] \times [d, d],$$
$$[a, a] \times [c, d], \qquad [b, b] \times [c, d].$$

Las mismas cuatro celdas también son la frontera de la celda $(a, b) \times (c, d)$.] Se infiere que una celda en $\boldsymbol{R}^p$ tiene contenido; más aún, fácilmente se puede ver que si $J = [a_1, b_1] \times \cdots \times [a_p, b_p]$, entonces

$$c(J) = \int_J 1 = (b_1 - a_1) \cdots (b_p - a_p).$$

Por lo tanto, el contenido de una celda, como se da en la definición 44.2, es consistente con la definición del contenido asignado a una celda cerrada en la sección 43. Observaciones análogas son aplicables a otras celdas en $\boldsymbol{R}^p$; en particular, se puede ver que si $K = [a_1, b_1) \times \cdots \times [a_p, b_p)$, entonces

$$c(K) = \int_K 1 = (b_1 - a_1) \cdots (b_p - a_p).$$

Se demostrará en seguida que el concepto de contenido cero que se introdujo en la definición 43.1 es consistente con el concepto de contenido que se introdujo en la definición 44.2.

44.3 LEMA. *Un conjunto* $A \subseteq \boldsymbol{R}^p$ *tiene contenido cero en el sentido de la definición 43.1 si y sólo si tiene contenido (en el sentido de la definición 44.2) y* $c(A) = 0$.

DEMOSTRACION. Supóngase que $A \subseteq \boldsymbol{R}^p$ tiene contenido cero. Entonces, si $\varepsilon > 0$, se puede encerrar a A en la unión U de un número finito de celdas cerradas con contenido total menor a ε. Dado que esta unión U es un conjunto acotado, entonces A es acotado; dado que U es cerrado, también contiene a $b(A)$. Como $\varepsilon > 0$ es arbitraria, se infiere que $b(A)$ tiene contenido cero; por lo tanto, A tiene contenido y

$$c(A) = \int_A 1.$$

Del lema 43.7 se sigue ahora que $c(A) = 0$. Inversamente, supóngase que $A \subseteq \boldsymbol{R}^p$ tiene contenido y que $c(A) = 0$. Por lo que hay una celda cerrada

I que contiene a A y tal que la función

$$g_I(x) = 1 \quad \text{para} \quad x \in A,$$
$$= 0 \quad \text{para} \quad x \in I \setminus A,$$

es integrable en I. Sea $\varepsilon > 0$ dada y sea P_ε una partición de I tal que cualquier suma de Riemann correspondientes a P_ε satisfaga $0 \leq S(P_\varepsilon; g_I) < \varepsilon$. Si se toman los puntos intermedios en $S(P_\varepsilon; g_I)$ como pertenecientes a A, siempre que sea posible, se infiere que A está contenido en la unión de un número finito de celdas en P con contenido total menor a ε. De modo que A tiene contenido cero en el sentido de la definición 43.1. Q.E.D.

44.4 TEOREMA. Supóngase que A, B pertenecen a $\mathcal{D}(\mathbf{R}^p)$ y sea $x \in \mathbf{R}^p$.

(a) Los conjuntos $A \cap B$ y $A \cup B$ pertenecen a $\mathcal{D}(\mathbf{R}^p)$ y

$$c(A) + c(B) = c(A \cap B) + c(A \cup B).$$

(b) Los conjuntos $A \setminus B$ y $B \setminus A$ pertenecen a $\mathcal{D}(\mathbf{R}^p)$ y

$$c(A \cup B) = c(A \setminus B) + c(A \cap B) + c(B \setminus A).$$

(c) Si $x + A = \{x + a : a \in A\}$, entonces $x + A$ pertenece a $\mathcal{D}(\mathbf{R}^p)$ y

$$c(x + A) = c(A).$$

DEMOSTRACION. Por hipótesis, las fronteras $b(A)$, $b(B)$ tienen contenido cero. Se deja como ejercicio para el lector demostrar que las fronteras

$$b(A \cap B), \quad b(A \cup B), \quad b(A \setminus B), \quad b(B \setminus A)$$

están contenidas en $b(A) \cup b(B)$, De esto y del ejemplo 43.2(c) se sigue que $A \cap B$, $A \cup B$, $A \setminus B$, y $B \setminus A$ pertenecen a $\mathcal{D}(\mathbf{R}^p)$.

Sea ahora I una celda cerrada que contiene a $A \cup B$ y sean f_a, f_b, f_i, f_u las funciones igual a 1 en A, B, $A \cap B$, $A \cup B$,espectivamente, e igual a 0 en las demás partes en I. Dado que cada una de estas funciones es continua excepto en conjuntos con contenido cero, son integrables en I Como

$$f_a + f_b = f_i + f_u,$$

del teorema 43.5 y de la definición de contenido se infiere que

$$c(A) + c(B) = \int_I f_a + \int_I f_b = \int_I (f_a + f_b)$$
$$= \int_I (f_i + f_u) = \int_I f_i + \int_I f_u$$

$$= c(A \cap B) + c(A \cup B).$$

Esto prueba la fórmula dada en *(a)*, la de *(b)* se puede probar de manera análoga.

Para probar *(c)*, obsérvese que si $\varepsilon > 0$ está dada y si $J_1, \ldots, J_n$ son celdas con contenido total menor a ε cuya unión contiene a $b(A)$, entonces $x + J_1, \ldots, x + J_n$ son celdas con contenido total menor a ε cuya unión contiene a $b(x + A)$. Dado que $\varepsilon > 0$ es arbitraria, el conjunto $x + A$ Para pertenece a $\mathcal{D}(\mathbf{R}^p)$. probar que $c(x + A) = c(A)$, sea I una celda cerrada que contiene a A, por lo que $x + I$ es una celda cerrada que contiene a $x + A$. Sea $f_1 : I \to \mathbf{R}$ tal que $f_1(y) = 1$ para $y \in A$ y $f_1(y) = 0$ para $y \in I \setminus A$, y sea $f_2 : x + I \to \mathbf{R}$ tal que $f_2(z) = 1$ para $z \in x + A$ y $f_2(z) = 0$ para $z \in (x + I) \setminus (x + A)$. Demostrar que a cada suma de Riemann para f_1 corresponde una suma de Riemann para f_2 y son iguales. Por lo tanto,

$$c(A) = \int_I f_1 = \int_{x+I} f_2 = c(x + A). \qquad \text{Q.E.D.}$$

44.5 COROLARIO. *Supóngase que A y B pertenecen a $\mathcal{D}(\mathbf{R}^p)$. (a) Si $A \cap B = \emptyset$, entonces $c(A \cup B) = c(A) + c(B)$. (b) Si $A \subseteq B$, entonces $c(B \setminus A) = c(B) - c(A)$.*

Caracterización de la función contenido

Ya se ha visto que la función contenido $c : \mathcal{D}(\mathbf{R}^p) \to \mathbf{R}$ es no negativa, "aditiva", "invariante bajo la traslación" y asigna el valor 1 al cubo "semi-abierto"

$$K_0 = [0, 1) \times [0, 1) \times \cdots \times [0, 1).$$

Se habrá de probar ahora que estas cuatro propiedades caracterizan a c.

44.6 TEOREMA. *Sea $\gamma : \mathcal{D}(\mathbf{R}^p) \to \mathbf{R}$ una función con las siguientes propiedades:*

(i) $\gamma(A) \geq 0$ *para toda* $A \in \mathcal{D}(\mathbf{R}^p)$;

(ii) si $A, B \in \mathcal{D}(\mathbf{R}^p)$ *y* $A \cap B = \emptyset$, *entonces* $\gamma(A \cup B) = \gamma(A) + \gamma(B)$;

(iii) si $A \in \mathcal{D}(\mathbf{R}^p)$ *y* $x \in \mathbf{R}^p$, *entonces* $\gamma(A) = \gamma(x + A)$;

(iv) $\gamma(K_0) = 1$.

Entonces, se tiene $\gamma(A) = c(A)$ *para toda* $A \in \mathcal{D}(\mathbf{R}^p)$.

DEMOSTRACION. Si $n \in \mathbf{N}$, sea K_n el cubo "semi-abierto"

$$K_n = [0, 2^{-n}) \times [0, 2^{-n}) \times \cdots \times [0, 2^{-n}).$$

Obsérvese que K_0 es la unión de 2^{np} traslaciones ajenas de K_n; por lo tanto, $1 = \gamma(K_0) = 2^{np}\gamma(K_n)$ y entonces $\gamma(K_n) = 1/2^{np} = c(K_n)$. Supóngase que A, B pertenecen a $\mathcal{D}(\mathbf{R}^p)$ y sea $A \subseteq B$. Entonces se puede escribir $B = A \cup (B \setminus A)$ ya que $A \cap (B \setminus A) = \emptyset$, de *(i)* y *(ii)* se infiere que

$$\gamma(B) = \gamma(A) + \gamma(B \setminus A) \geq \gamma(A).$$

Por lo tanto, γ es monótona en el sentido de que si $A \subseteq B$ entonces $\gamma(A) \leq \gamma(B)$. Sea ahora $A \in \mathcal{D}(\mathbf{R}^p)$. Dado que A es acotado, entonces, para alguna $M \in \mathbf{N}$, el conjunto A está contenido en el interior del cubo cerrado I con longitud media 2^M y centro en 0. Si $\varepsilon > 0$, hay una partición de I en cubos pequeños de longitud lateral 2^{-n}, digamos, tal que el contenido de la unión de todos los cubos $I_1, \ldots, I_r$ que están contenidos en A exceda $c(A) - \varepsilon$, y tal que el contenido de todas las celdas $I_1, \ldots, I_s$ $(r \leq s)$ que contengan puntos de A no exceda a $c(A) + \varepsilon$. (Véase el ejercicio 44.I.) Ahora, cada uno de estos conjuntos I_i difiere de una traslación $x_i + K_n$ en un conjunto de contenido cero. Por lo tanto, se tiene

$$c(A) - \varepsilon \leq c\left(\bigcup_{i=1}^{r} (x_i + K_n)\right) \leq c(A) \leq c\left(\bigcup_{i=1}^{s} (x_i + K_n)\right) \leq c(A) + \varepsilon.$$

Ahora, c y γ son ambas invariantes bajo la traslación del conjunto y coinciden en K_n. Más aún, c y γ son aditivas sobre uniones ajenas finitas. Por lo tanto, se sigue que

$$c\left(\bigcup_{i=1}^{r} (x_i + K_n)\right) = \sum_{i=1}^{r} c(x_i + K_n) = \sum_{i=1}^{r} \gamma(x_i + K_n)$$
$$= \gamma\left(\bigcup_{i=1}^{r} (x_i + K_n)\right).$$

De esto y del hecho de que γ es monótona se infiere que

$$c(A) - \varepsilon \leq \gamma\left(\bigcup_{i=1}^{r} (x_i + K_n)\right) \leq \gamma(A) \leq \gamma\left(\bigcup_{i=1}^{s} (x_i + K_n)\right) \leq c(A) + \varepsilon,$$

de donde $|\gamma(A) - c(A)| \leq \varepsilon$. Como $\varepsilon > 0$ es arbitraria, se infiere que $\gamma(A) = c(A)$.

Q.E.D.

44.7 COROLARIO. *Sea $\mu : \mathcal{D}(\mathbf{R}^p) \to \mathbf{R}$ una función que satisface las propiedades (i), (ii) y (iii). Entonces, existe una constante $m \geq 0$ tal que $\mu(A) = mc(A)$ para toda $A \in \mathcal{D}(\mathbf{R}^p)$.*

DEMOSTRACION. Dado que μ posee las propiedades *(i)* y *(ii)*, con facilidad se puede ver que μ es monótona en el sentido de que $A \subseteq B$ implica que $\mu(A) \leq \mu(B)$. Si $\mu(K_0) = 0$, entonces μ de cualesquier conjuntos acota-

dos es 0, de donde se sigue que $\mu(A) = 0$ para toda $A \in \mathcal{D}(\mathbf{R}^p)$, de modo que se puede tomar $m = 0$. Si $\mu(K_0) \neq 0$, sea

$$\gamma(A) = \frac{1}{\mu(K_0)}\,\mu(A) \qquad \text{para toda} \quad A \in \mathcal{D}(\mathbf{R}^p).$$

Puesto que con facilidad se puede ver que γ tiene las propiedades *(i)*, *(ii)*, *(iii)* y *(iv)* del teorema, se infiere que $\gamma = c$. Por lo tanto, se toma $m = \mu(K_0)$.

Otras propiedades de la integral

Se ofrecerán ahora algunas propiedades adicionales de la integral que a menudo son útiles.

44.8 TEOREMA. *Sea* $A \in \mathcal{D}(\mathbf{R}^p)$ *y supóngase que* $f: A \to \mathbf{R}$ *es acotada y continua en* A. *Entonces,* f *es integrable en* A.

DEMOSTRACION. Sea I una celda cerrada con $A \subseteq I$ y supóngase que $f_I : I \to \mathbf{R}$ es igual a f en A e igual a 0 en $I \setminus A$. Dado que f_I es acotada en I y es continua en $I \setminus b(A)$, del teorema de integrabilidad 43.9 se infiere que f_I es integrable en I. Por lo tanto, f es integrable en A.

Se probará en seguida que la integral es aditiva con respecto al conjunto sobre el que se extiende la integral.

44.9 TEOREMA. *(a) Supóngase que* A_1 *y* A_2 *pertenecen a* $\mathcal{D}(\mathbf{R}^p)$ *y supóngase que* $A_1 \cap A_2$ *tiene contenido cero. Si* $A = A_1 \cup A_2$ *y si* $f: A \to \mathbf{R}$ *es integrable en* A_1 *y* A_2, *entonces* f *es integrable en* A *y*

$$\int_A f = \int_{A_1} f + \int_{A_2} f. \tag{44.1}$$

(b) Supóngase que A *pertenece a* $\mathcal{D}(\mathbf{R}^p)$ *y sean* A_1, $A_2 \in \mathcal{D}(\mathbf{R}^p)$ *tales que* $A = A_1 \cup A_2$ *y tales que* $A_1 \cap A_2$ *tiene contenido cero. Si* $f: A \to \mathbf{R}$ *es integrable en* A *y si las restricciones de* f *a* A_1 *y* A_2 *son integrables, entonces (44.1) es válido*

DEMOSTRACION. *(a)* Sea I una celda cerrada que contiene a $A = A_1 \cup A_2$ y supóngase que $f_i : I \to \mathbf{R}$, $i = 1, 2$, es igual a f en A_i e igual a 0 en cualquier otra parte en I. Por hipótesis, f_1 y f_2 son integrables en I y

$$\int_I f_i = \int_{A_i} f, \qquad i = 1, 2.$$

Del teorema 43.5 se sigue que $f_1 + f_2$ es integrable en I y que

$$\int_I (f_1 + f_2) = \int_I f_1 + \int_I f_2.$$

Ahora, como $f(x) = f_1(x) + f_2(x)$ siempre que $x \in A \setminus (A_1 \cap A_2)$, del lema 43.8 se deduce que f es integrable en A y que (44.1) es válido.

(b) Se preservará la notación de la demostración de *(a)*. Por hipótesis, f_I es integrable en I. Ahora, $f_I(x) = f_1(x) + f_2(x)$ excepto para x en $A_1 \cap A_2$, un conjunto con contenido cero. Por lo tanto, del teorema 43.5 y el lema 43.8 se infiere que

$$\int_A f = \int_I f = \int_I (f_1 + f_2) = \int_I f_1 + \int_I f_2$$

$$= \int_{A_1} f + \int_{A_2} f. \qquad \text{Q.E.D.}$$

Se hace la aclaración de que si $f: A \to \mathbf{R}$s una función acotada integrable, entonces se satisface automáticamente el supuesto que se hizo en 44.9*(b)* de que las restricciones de f a A_1 y A_2 son integrables. (Véase el ejercicio 44.J.)

El siguiente resultado a menudo es útil para calcular la magnitud de una integral.

44.10 TEOREMA. *Sea $A \in \mathfrak{D}(\mathbf{R}^p)$ y supóngase que $f: A \to \mathbf{R}$ es integrable en A y es tal que $|f(x)| \le M$ para toda $x \in A$. Entonces*

(44.2) $$\left| \int_A f \right| \le Mc(A).$$

De manera más general, si f es de valor real y $m \le f(x) \le M$ para toda $x \in A$, entonces

(44.3) $$mc(A) \le \int_A f \le Mc(A).$$

DEMOSTRACION. Sea f_I la extensión de f a una celda cerrada I que contiene a A. Si $\varepsilon > 0$ está dada, entonces existe una partición $P_\varepsilon = \{J_1, \ldots, J_h\}$ de I tal que si $S(P_\varepsilon; f_I)$ es cualquier suma de Riemann correspondiente, entonces

$$S(P_\varepsilon; f_I) - \varepsilon \le \int_I f_I \le S(P_\varepsilon; f) + \varepsilon.$$

Obsérvese que si los puntos intermedios de la suma de Riemann se eligen fuera de A, cuando sea posible, se tiene

$$S(P_\varepsilon; f) = \sum{}' f(x_j) c(J_j),$$

en donde la suma se extiende sobre aquellas celdas en P_ε completamente contenidas en A. De donde

$$S(P_\varepsilon; f) \le M \sum{}' c(J_k) \le Mc(A).$$

i465 (1689x2581x2 pcx)

Por lo tanto, se tiene

$$\int_A f = \int_I f_I \leq Mc(A) + \varepsilon,$$

y puesto que $\varepsilon > 0$ es arbitraria se obtiene el lado derecho de la desigualdad (44.3). El lado izquierdo se prueba de manera análoga.

Como consecuencia de este resultado se obtiene el siguiente teorema, que es una extensión del primer teorema del valor medio 30.6.

44.11 TEOREMA DEL VALOR MEDIO. *Sea* $A \in \mathcal{D}(\boldsymbol{R}^p)$ *conjunto conexo y supóngase que* $f: A \to \boldsymbol{R}$ *es acotada y continua en* A. *Entonces, existe un punto* $p \in A$ *tal que*

$$\int_A f = f(p)c(A). \tag{44.4}$$

DEMOSTRACION. Si $c(A) = 0$, la conclusión es trivial. Por lo tanto, se va a suponer que $c(A) \neq 0$. Sea

$$m = \inf\{f(x): x \in A\}, \qquad M = \sup\{f(x): x \in A\};$$

de la segunda parte del teorema anterior se sigue que

$$m \leq \frac{1}{c(A)} \int_A f \leq M. \tag{44.5}$$

Si ambas desigualdades en (44.5) son estrictas, el resultado se deduce del teorema del valor intermedio de Bolzano 22.4.

Ahora, supóngase que $\int_A f = Mc(A)$. Si el supremo M se alcanza en algún $p \in A$, la conclusión también se sigue. Por ello, se va suponer que el supremo M no se alcanza en A. Dado que $c(A) \neq 0$, existe una celda cerrada $K \subseteq A$ tal que $c(K) \neq 0$ (véase el ejercicio 44.G). Como K es compacto y f es continua en K, existe $\varepsilon > 0$ tal que $f(x) \leq M - \varepsilon$ para toda $x \in K$. Dado que $A = K \cup (A \setminus K)$ de los teoremas 44.9 y 44.10 se infiere que

$$\begin{aligned} Mc(A) = \int_A f &= \int_K f + \int_{A \setminus K} f \\ &\leq (M - \varepsilon)c(K) + Mc(A \setminus K) < Mc(A), \end{aligned}$$

lo que es una contradicción. Si $\int_A f = mc(A)$, entonces un argumento análogo es aplicable.

La integral como una integral iterada

Es conveniente saber que si f es integrable en una celda cerrada $J = [a_1, b_1] \times \cdots \times [a_p, b_p]$ y $\boldsymbol{R}^p$ y tiene valores en $\boldsymbol{R}$, entonces la integral $\int_J f$

se puede calcular en términos de una "integral iterada"

$$\int_{a_p}^{b_p}\left\{\cdots\left\{\int_{a_2}^{b_2}\left\{\int_{a_1}^{b_1} f(x_1, x_2, \ldots, x_p)\, dx_1\right\} dx_2\right\}\cdots\right\} dx_p.$$

Este es el método para calcular integrales dobles y triples por medio de integrales iteradas que le es familiar al lector desde el cálculo elemental. Se dará una justificación de este procedimiento. Para simplificar se va a suponer que $p=2$, pero claramente el resultado se extiende a dimensiones más altas.

44.12. TEOREMA. *Si f es continua de una celda cerrada* $J=[a, b]\times[c, d]$ a $\mathbf{R}$, entonces

$$\int_J f = \int_c^d\left\{\int_a^b f(x, y)\, dx\right\} dy$$
$$= \int_a^b\left\{\int_c^d f(x, y)\, dy\right\} dx.$$

DEMOSTRACION. En el teorema de intercambio 31.9 se vio que las dos integrales iteradas son iguales. Para probar que la integral de f en J está dada por la primera integral iterada, defínase F para $y\in[c, d]$ como

$$F(y)=\int_a^b f(x, y)\, dx.$$

Sea $c=y_0\le y_1\le\cdots\le y_r=d$ una partición del intervalo $[c, d]$, sea $a=x_0\le x_1\le\cdots\le x_s=b$ una partición de $[a, b]$, y denótese por P a la partición de J que se obtiene usando las celdas $[x_{k-1}, x_k]\times[y_{j-1}, y_j]$. Sea y_j^* cualquier punto en $[y_{j-1}, y_j]$ y obsérvese que

$$F(y_j^*)=\int_a^b f(x, y_j^*)\, dx=\sum_{k=1}^{s}\int_{x_{k-1}}^{x_k} f(x, y_j^*)\, dx.$$

De acuerdo con el primer teorema del valor medio 30.6, para cada j y k existe un punto x_{jk}^* en $[x_{k-1}, x_k]$ tal que

$$F(y_j^*)=\sum_{k=1}^{s} f(x_{jk}^*, y_j^*)(x_k-x_{k-1}).$$

Se multiplica por (y_j-y_{j-1}) y se suma para obtener

$$\sum_{j=1}^{r} F(y_j^*)(y_j-y_{j-1})=\sum_{j=1}^{r}\sum_{k=1}^{s} f(x_{jk}^*, y_j^*)(x_k-x_{k-1})(y_j-y_{j-1}).$$

Ahora, la expresión del lado izquierdo de esta fórmula es una suma arbitraria de Riemann para la integral

$$\int_c^d F(y)\,dy = \int_c^d \left\{ \int_a^b f(x, y)\,dx \right\} dy.$$

Se ha probado que esta suma de Riemann es igual a una suma de Riemann particular correspondiente a la partición P. Dado que f es integrable en J, la existencia de esta integral iterada y su igualdad con la integral sobre J está probada. Q.E.D.

Una modificación mínima de la demostración dada para el teorema anterior da origen a la siguiente aseveración un poco más fuerte.

44.13 TEOREMA. *Sea f integrable del rectángulo $J = [a, b] \times [c, d]$ $\mathbf{R}$ y supóngase que, para cada $y \in [c, d]$, la función $x \mapsto f(x, y)$ de $[a, b]$ en $\mathbf{R}$ es continua excepto posiblemente para un número finito de puntos, en los que tiene límites unilaterales. Entonces*

$$\int_J f = \int_c^d \left\{ \int_a^b f(x, y)\,dx \right\} dy.$$

Como consecuencia de este teorema, se obtendrá un resultado que a menudo se usa para calcular integrales de funciones definidas en un conjunto acotado por curvas continuas. Por conveniencia, se establecerá el resultado para el caso en que el conjunto tiene segmentos de línea horizontales como frontera superior e inferior y curvas continuas como sus fronteras laterales. Es claro que, un resultado análogo es válido si las fronteras laterales son segmentos de líneas verticales y las fronteras superior e inferior son curvas. Conjuntos más complicados se tratan descomponiéndolos en la unión de subconjuntos de estos dos tipos.

44.14 TEOREMA. *Suponga que $A \subseteq \mathbf{R}^2$ está dado por*

$$A = \{(x, y) : \alpha(y) \le x \le \beta(y),\ c \le y \le d\},$$

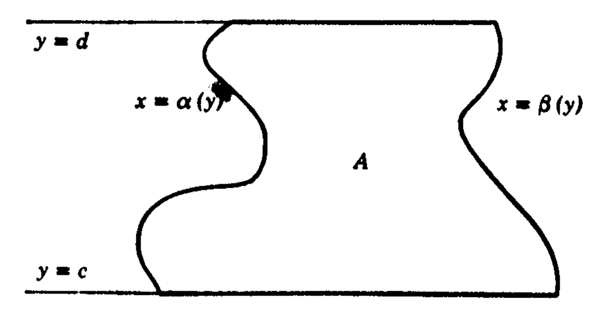


Figura 44.1

en donde α y β son funciones continuas en $[c, d]$ con valores en $[a, b]$. Si f es continua en $A \to \mathbf{R}$, entonces f es integrable en A y

$$\int_A f = \int_c^d \left\{ \int_{\alpha(y)}^{\beta(y)} f(x, y)\, dx \right\} dy.$$

DEMOSTRACION. Sea J una celda cerrada que contiene a A y sea f_J la extensión de f a J. Una variación del ejemplo 43.2(*f*) prueba que la frontera de A tiene contenido cero, por lo que f_J es integrable en J. Ahora, para cada $y \in [c, d]$ la función $x \mapsto f_J(x, y)$ es continua excepto posiblemente en los dos puntos $\alpha(y)$ $\beta(y)$,en los que tiene límites unilaterales. Del teorema anterior se imfiere que

$$\int_A f = \int_J f_J = \int_c^d \left\{ \int_a^b f_J(x, y)\, dx \right\} dy$$
$$= \int_c^d \left\{ \int_{\alpha(y)}^{\beta(y)} f(x, y)\, dx \right\} dy. \qquad \text{Q.E.D.}$$

Ejercicios

44.A. Si $A \subseteq \boldsymbol{R}^p$, un punto es un punto frontera de A si y sólo si es un punto frontera del complemento $\mathscr{C}(A)$ de A. De donde, $b(A) = b(\mathscr{C}(A))$.

44.B. Sea $A \subseteq \boldsymbol{R}^p$, y sea $b(A)$ la frontera de A.

(a) El conjunto $b(A)$ es cerrado en $\boldsymbol{R}^p$.

(b) El interior $A^0 = A \setminus b(A)$es abierto en$\boldsymbol{R}^p$ y contiene a todo conunto abierto G tal que $G \subseteq A$.

(c) La cerradura $A^- = A \cup b(A)$ es cerrada en $\boldsymbol{R}^p$ y está contenida en todo conjunto cerrado F tal que $A \subseteq F$.

44.C. Sea $A \subseteq \boldsymbol{R}^p$ y sea $A^- = A \cup b(A)$ la cerradura de A. Demostrar que $b(A^-) \subseteq b(A)$. Dar un ejemplo para probar que la igualdad se puede cumplir y un ejemplo de que la igualdad puede no cumplirse.

44.D. Sean A, B subconjunto de $\boldsymbol{R}^p$. Demostrar que la frontera de cada uno de los conjuntos

$$A \cap B, \qquad A \setminus B, \qquad A \cup B$$

está contenida en $b(A) \cup b(B)$. *(Sugerencia:* $b(A) = A^- \cap (\mathscr{C}(A))^-$.)

44.E. Un conjunto $A \subseteq \boldsymbol{R}^p$ es cerrado en $\boldsymbol{R}^p$ si y sólo si $b(A) \subseteq A$. Un conjunto $B \subseteq \boldsymbol{R}^p$ es abierto en $\boldsymbol{R}^p$ si y sólo si $B \cap b(B) = \emptyset$.

44.F. Si $A \in \mathscr{D}(\boldsymbol{R}^p)$, demostrar que su interior $A^0 = A \setminus b(A)$ y su cerradura $A^- = A \cup b(A)$ también pertenecen a $\mathscr{D}(\boldsymbol{R}^p)$ y que $c(A^0) = c(A) = c(A^-)$.

44.G. Si $A \in \mathscr{D}(\boldsymbol{R}^p)$ y $c(A) > 0$, demostrar que existe una celda cerrada $K \subseteq A$ tal que $c(K) \neq 0$.

44.H. Si $A \subseteq \boldsymbol{R}^p$, se define el **contenido interno** y **externo** de A como

$$c_*(A) = \sup c(U), \qquad c^*(A) = \inf c(V),$$

en donde el supremo se toma sobre el conjunto de todas las uniones finitas de celdas contenidas en A y el ínfimo se toma sobre el conjunto de todas las uniones finitas de celdas que contienen puntos de A.

(a) Demostrar que $c_*(A) \leq c^*(A)$ y que A tiene contenido si y sólo si $c_*(A) = c^*(A)$, en cuyo caso $c(A)$ es este valor común.

(*b*) Si A y B son subconjuntos ajenos de $\mathbf{R}^p$, demostrar que $c^*(A \cup B) \leq c^*(A) + c^*(B)$.

(*c*) Dar un ejemplo de conjuntos ajenos A, B tales que $0 \neq c^*(A) = c^*(B) = c^*(A \cup B)$.

44.I. Sea $M \in \mathbf{N}$ y sea $I_M \subseteq \mathbf{R}^p$ el cubo con longitud media 2^M y centro 0. Para $n \in \mathbf{N}$ se divide I_M en una **red** $G_{M,n}$ de longitud 2^{-n} formada por la colección de todos los cubos en I_M con longitud lateral 2^{-n} y puntos extremos racionales bivalentes (es decir, puntos extremos de la forma $k/2^n$ en donde $k \in \mathbf{Z}$).

(*a*) Si $J \subseteq I_M$ es una celda cerrada y $\varepsilon > 0$, demostrar que existe $n \in \mathbf{N}$ tal que la unión de todos los cubos en $G_{M,n}$ que están contenidos en J tiene contenido total mayor que $c(J) - \varepsilon$ y la unión de todos los cubos en $G_{M,n}$ que contienen puntos en J tiene contenido total menor a $c(J) + \varepsilon$.

(*b*) Si $A \subseteq I_M$ tiene contenido y $\varepsilon > 0$, demostrar que existe $n \in \mathbf{N}$ tal que la unión de todos los cubos en $G_{M,n}$ que están contenidos en A tiene contenido total que excede a $c(A) - \varepsilon$ y la unión de todos los cubos en $G_{M,n}$ que contienen puntos en A tiene contenido total menor a $c(A) + \varepsilon$.

44.J. Sea $I \subseteq \mathbf{R}^p$ una celda cerrada y sea $f: I \to \mathbf{R}$ integrable en I. Si $A \subseteq I$ tiene contenido, entonces la restricción de f a A es integrable en A. (*Sugerencia* usar el ejercicio 43.P.)

44.K. Sea $A \in \mathcal{D}(\mathbf{R}^p)$ y supóngase que f y g son integrables en A y que $g(x) \geq 0$ para toda $x \in A$. Si $m = \inf f(A)$, $M = \sup f(A)$, entonces existe un número real $\mu \in [m, M]$ tal que

$$\int_A fg = \mu \int_A g.$$

44.L. Si además de las hipótesis del ejercicio anterior se supone que A es conexo y f es continua en A, entonces existe un punto $p \in A$ tal que

$$\int_A fg = f(a) \int_A g.$$

44.M. Sea $\{(x_n, y_n): n \in \mathbf{N}\}$ una enumeración de los puntos en $(0, 1) \times (0, 1)$ con coordenadas racionales. Para cada $n \in \mathbf{N}$, sea I_n una celda abierta en $(0, 1) \times (0, 1)$ que contiene a (x_n, y_n) y sea $G = \bigcup_{n \in \mathbf{N}} I_n$. Demostrar que G es un conjunto abierto en $\mathbf{R}^2$ cuya frontera $b(G)$ es $(0, 1) \times (0, 1) \setminus G$. Demostrar que si $\sum c(I_n) < 1$, entonces el conjunto abierto G no tiene contenido.

44.N. Usando la terminología del ejercicio 7.K, sea $A \subseteq [0, 1]$ un conjunto "estilo Cantor" con longitud $\frac{1}{2}$. Si $K = A \times [0, 1]$, demostrar que K es un subconjunto compacto de $\mathbf{R}^2$, que $b(K) = K$ y que K no tiene contenido.

44.O. Sea $a \leq b$ y supongase que $f: [a, b] \to \mathbf{R}$ es continua y tal que $f(x) \geq 0$ para toda $x \in [a, b]$. A $S_f = \{(x, y): a \leq x \leq b,\ 0 \leq y \leq f(x)\}$ se le da el nombre de conjunto ordenado de f. Examinando la frontera de S_f, demostrar que tiene contenido. Probar que

$$c(S_f) = \int_a^b \left\{ \int_0^{y=f(x)} 1 dy \right\} dx = \int_a^b f(x)\, dx.$$

44.P. Sea $A \subseteq \mathbf{R}^2$ el conjunto del ejercicio 43.E y supóngase que f está definida en $Q = [0, 1] \times [0, 1] \to \mathbf{R}$ como $f(x, y) = 1$ para $(x, y) \in A$ y $f(x, y) = 0$ de lo contrario.

Demostrar que A no tiene contenido y que f no es integrable en Q. Sin embargo, las integrales iteradas existen y satisfacen

$$\int_0^1 \left\{ \int_0^1 f(x, y)\, dx \right\} dy = \int_0^1 \left\{ \int_0^1 f(x, y)\, dy \right\} dx.$$

44.Q. Sea $Q = [0, 1] \times [0, 1]$ y supóngase que $f: Q \to \boldsymbol{R}$ está definida como $f(x, y) = 0$ si x o y es irracional y $f(x, y) = 1/n$ si y es racional y $x = m/n$ en donde m y $n > 0$ son enteros primos relativos. Demostrar que

$$\int_Q f = \int_0^1 \left\{ \int_0^1 f(x, y)\, dx \right\} dy = 0,$$

pero que $\int_0^1 f(x, y)\, dy$ no existe para x racional.

44.R. Sea $J \subseteq \boldsymbol{R}^2$ una celda abierta que contiene a (0,0) y sea $f: J \to \boldsymbol{R}$ continua en J. Defínase $F: J \to \boldsymbol{R}$ por la integral iterada:

$$F(x, y) = \int_0^x \left\{ \int_0^y f(s, t)\, dt \right\} ds.$$

Demostrar que $D_2 D_1 F(x, y) = f(x, y) = D_1 D_2 F(x, y)$ para $(x, y) \in J$.

44.S. Sea J como en el ejercicio anterior y sea $G: J \to \boldsymbol{R}$ tal que $D_2 D_1 G$ es continua en J. Usar este ejercicio para probar que $D_1 D_2 G$ existe y es igual a $D_2 D_1 G$.

44.T. Sea $J = [a_1, b_1] \times \cdots \times [a_p, b_p]$ y sea $f: J \to \boldsymbol{R}$ continua. Sea $J_{(1)} = [a_2, b_2] \times \cdots \times [a_p, b_p]$ en $\boldsymbol{R}^{p-1}$ y defínase $F_{(1)}: J_{(1)} \to \boldsymbol{R}$ como

$$F_{(1)}(x_2, \ldots, x_p) = \int_{a_1}^{b_1} f(x_1, x_2, \ldots, x_p)\, dx_1.$$

(a) Demostrar que $F_{(1)}$ es continua en $J_{(1)}$.

(b) Dado $(x_2^*, \ldots, x_p^*)$ en $J_{(1)}$ y cualquier partición $a_1 = x_{1,0} < x_{1,1} < \cdots < x_{1,n} = b_1$ de $[a_1, b_1]$, demostrar que existen puntos $x_{1,k}^*$ en $[x_{1,k-1}, x_{1,k}]$ tales que

$$F_{(1)}(x_2^*, \ldots, x_p^*) = \sum_{k=1}^{n} f(x_{1,k}^*, x_2^*, \ldots, x_p^*)(x_{1,k} - x_{1,k-1}).$$

(c) Demostrar que

$$\int_{J_{(1)}} F_{(1)}(x_2, \ldots, x_p)\, d(x_2, \ldots, x_p) = \int_{J_{(1)}} \left\{ \int_{a_1}^{b_1} f(x_1, x_2, \ldots, x_p)\, dx_1 \right\} d(x_2, \ldots, x_p)$$
$$= \int_J f(x_1, \ldots, x_p)\, d(x_1, \ldots, x_p).$$

(d) Extender el resultado para el caso en que para cada punto $(x_2, \ldots, x_p)$ en $J_{(1)}$ la función $x_1 \mapsto F(x_1, x_2, \ldots, x_p)$ de $[a_1, b_1] \to \boldsymbol{R}$ es continua excepto posiblemente para un número finito de puntos en los que tiene límites unilaterales.

44.U. *(a)* Sean $\alpha, \beta: [a, b] \to \boldsymbol{R}$ continuas con $\alpha(x) \leq \beta(x)$ para toda $x \in [a, b]$. Demostrar que el conjunto

$$B = \{(x, y) \in \boldsymbol{R}^2 : a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}$$

es un conjunto compacto en $\boldsymbol{R}^2$ con contenido.

(b) Ahora, sean γ, $\delta: B \to \boldsymbol{R}$ funciones continuas con $\gamma(x, y) \leq \delta(x, y)$ para toda $(x, y) \in B$. Demostrar que el conjunto

$$D = \{(x, y, z) \in \boldsymbol{R}^3 : (x, y) \in B, \gamma(x, y) \leq z \leq \delta(x, y)\}$$

es un conjunto compacto en $\boldsymbol{R}^3$ con contenido.

(c) Si $f: D \to \boldsymbol{R}$ es continua, demostrar que f es integrable en D y que

$$\int_D f = \int_a^b \left\{ \int_{\alpha(x)}^{\beta(x)} \left\{ \int_{\gamma(x,y)}^{\delta(x,y)} f(x, y, z)\, dz \right\} dy \right\} dx.$$

44.V. Sea $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$ y para cada $j = 1, \ldots, p$, supóngase que $f_j: [a_j, b_j] \to \boldsymbol{R}$ es una función integrable. Si $\varphi: I \to \boldsymbol{R}$ está definida como $\varphi(x_1, \ldots, x_p) = f_1(x_1) \cdots f_p(x_p)$, demostrar que φ es integrable en I y que

$$\int_I \varphi = \left\{ \int_{a_1}^{b_1} f_1 \right\} \cdots \left\{ \int_{a_p}^{b_p} f_p \right\}.$$

44.W. Usar el teorema de aproximación de Weierstrass para probar que si $I = [a_1, b_1] \times \cdots \times [a_p, b_p]$ y si $g: I \to \boldsymbol{R}$ es continua entonces

$$\int_I g = \int_{a_1}^{b_1} \left\{ \int_{a_2}^{b_2} \cdots \left\{ \int_{a_p}^{b_p} g(x_1, x_2, \ldots, x_p)\, dx_p \right\} \cdots dx_2 \right\} dx_1.$$

44.X. Sea $\varphi: [0, +\infty) \to \boldsymbol{R}$ con $\varphi(0) = 0$ continua, no acotada y estrictamente creciente y supóngase que ψ es su función inversa. Por lo tanto, ψ es también continua y estrictamente creciente en $[0, +\infty)$.

(a) Si α, β son números positivos, comparar el área del intervalo $[0, \alpha] \times [0, \beta]$ con las áreas acotadas por los ejes coordenados y la gráfica de φ para obtener la desigualdad de Young:

$$\alpha\beta \leq \int_0^\alpha \varphi + \int_0^\beta \psi.$$

(b) Si $p \geq 1$ y $q \geq 1$ son tales que $(1/p) + (1/q) = 1$, y si $\varphi(x) = x^{1/q}$ y $\psi(x) = y^{q/p}$, usar la desigualdad de Young para probar la desigualdad

$$\alpha\beta \leq \alpha^p/p + \beta^q/q.$$

(c) Si a_i, b_i, $i = 1, \ldots, n$ son números reales y si $A = (|a_1|^p + \cdots + |a_n|^p)^{1/p}$ y $B = (|b_1|^q + \cdots + |b_n|^q)^{1/q}$, usar la desigualdad anterior para deducir la desigualdad de Hölder

$$\sum_{i=1}^n |a_i b_i| \leq AB,$$

que se obtuvo en el proyecto 8.β(b).

Proyectos

44.α. Sea $I \subseteq \boldsymbol{R}^p$ una celda cerrada y supóngase que $f: I \to \boldsymbol{R}$ es acotada. Para $\alpha > 0$, sea $D_\alpha = \{x \in I : \omega_f(x) \geq a\}$, en donde $\omega_f(x)$ designa la oscilación de f en x (véase el proyecto 23.α).

(a) Supóngase que D_α tiene contendio cero. Sea $P_\alpha = \{I_1, \ldots, I_n\}$ una partición de I tal que (*i*) cada punto de D_α está contenido en el interior de una de las celdas $I_1, \ldots, I_r$ $(r \leq n)$, (ii) $c(I_1) + \cdots + c(I_r) \leq \alpha/2 \|f\|_I$, y *(iii)* si $x, y \in K_j$ para $j = r+1, \ldots, n$, entonces $|f(x) - f(y)| < \alpha$. Si P es un refinamiento de P_α, probar que $|S(P; f) - S(P_\alpha; f)| \leq \alpha(c(I) + 1)$.

(b) Deducir que si D_α tiene contenido cero para cada $\alpha > 0$ entonces f es integrable en I.

(c) Supóngase que para alguna $\alpha > 0$, el contenido externo $c^*(D_\alpha) > 0$. Demostrar que para cualquier partición $P = \{J_1, \ldots, J_n\}$ de I se tiene

$$\sum_{j=1}^{n} (M_j - m_j) c(J_j) \geq \alpha c^*(D_\alpha).$$

Deducir que f no es integrable en I.

(d) Concluir que f es integrabla en I si y sólo si el conjunto D_α tiene contenido cero para toda $\alpha > 0$.

(e) Recuérdese que $D = \bigcup_{n \in N} D_{1/n}$ es el conjunto de puntos en donde f es discontinua. Demostrar que D tiene medida cero (en el sentido del ejercicio 43.V) si y sólo si cada conjunto $D_{1/n}$ tiene contenido cero.

(f) Concluir que f es integrable en I si y sólo si su conjunto D de puntos de discontinuidad tiene medida cero. (Este resultado es el **criterio de Lebesgue para integrabilidad.**)

44.β. En este proyecto se consideran integrales inferiores y superiores (introducidas en el proyecto 43.α) y sus iteraciones. Sean $I \subseteq \mathbf{R}^r$ y $J \subseteq \mathbf{R}^s$ celdas cerradas, $p = r + s$, y sea $K = I \times J \subseteq \mathbf{R}^p = \mathbf{R}^r \times \mathbf{R}^s$. Supóngase que $f : K \to \mathbf{R}$ es acotada.

(a) Para cada $x \in I$, defínase $g_x : J \to \mathbf{R}$ como $\mu(x) = U(g_x)$ para $y \in J$. Defínase $\lambda : I \to \mathbf{R}$ como la integral inferior $g_x(y) = f(x, y)$ de g_x, y defínase $\mu : I \to \mathbf{R}$ como la integral superior $\lambda(x) = L(g_x)$ de g_x. Si R es cualquier partición de I y S es cualquier partición de J y $P = R \times S$ la partición resultante de K, entonces demostrar que

$$L(P; f) \leq L(R; \lambda) \leq U(R; \lambda) \leq U(R; \mu) \leq U(P; f).$$

(b) Probar que

$$L(f) \leq L(\lambda) \leq U(\lambda) \leq U(f), \qquad L(f) \leq L(\mu) \leq U(\mu) \leq U(f).$$

Por lo tanto, si f es integrable en K, entonces λ y μ son integrables en I y

$$\int_K f = \int_I \lambda = \int_I \mu.$$

(c) Para cada $y \in J$, defínase $h_y : I \to \mathbf{R}$ como $h_y(x) = f(x, y)$ para $x \in I$. Supóngse que $\lambda' : J \to \mathbf{R}$ y $\mu' : J \to \mathbf{R}$ están definidas como

$$\lambda'(y) = L(h_y), \qquad \mu'(y) = U(h_y).$$

Demostrar que si f es integrable en K, entonces λ' y μ' son integrables en J y

$$\int_K f = \int_J \lambda' = \int_J \mu'.$$

(d) Si $g_x : J \to \mathbf{R}$ es integrable en J para cada $x \in I$, entonces $\lambda = \mu$ y

$$\int_K f(x, y)\, d(x, y) = \int_K f = \int_J \left\{ \int_I f(x, y)\, dy \right\} dx.$$

Análogamente, si $h_y : I \to \boldsymbol{R}$ es integrable en I para cada $y \in J$, entonces

$$\int_K f(x, y)\, d(x, y) = \int_K f = \int_I \left\{ \int_J f(x, y)\, dx \right\} dy.$$

44.γ. Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y sea $\mathcal{D}(\Omega)$ la colección de todos los conjuntos $A \in \mathcal{D}(\boldsymbol{R}^p)$ con $A^- \subseteq \Omega$. En este proyecto se introducirá el concepto de una función "aditiva" en $\mathcal{D}(\Omega)$ y de su "fuerte densidad". Se dice que una función $G : \mathcal{D}(\Omega) \to \boldsymbol{R}$ es **aditiva** si

$$G(A \cup B) = G(A) + G(B)$$

siempre que A, $B \in \mathcal{D}(\Omega)$ y $A \cap B = \emptyset$.

(a) Si $f : \Omega \to \boldsymbol{R}$ es integrable en todo conjunto en $\mathcal{D}(\Omega)$ y si se define $F : \mathcal{D}(\Omega) \to \boldsymbol{R}$ como

$$F(A) = \int_A f,$$

entonces F es aditiva en $\mathcal{D}(\Omega)$.

(b) Sea $G : \mathcal{D}(\Omega) \to \boldsymbol{R}$ una función aditiva y sea $g : \Omega \to \boldsymbol{R}$. Se dice que g es una **densidad fuerte** para G si para toda $\varepsilon > 0$ y todo conjunto $A \in \mathcal{D}(\Omega)$, existe $\delta > 0$ tal que si K es un cubo cerrado con longitud lateral menor a δ contenido en Ω, y si $x \in A \cap K$, entonces

$$\left| \frac{G(K)}{c(K)} - g(x) \right| < \varepsilon.$$

(c) Sea $\Omega = \boldsymbol{R}^p$. Demostrar que la función contenido $c : \mathcal{D}(\boldsymbol{R}^p) \to \boldsymbol{R}$ tiene densidad fuerte idénticamente igual a 1 en $\boldsymbol{R}^p$.

(d) Supóngase que $\mu : \mathcal{D}(\boldsymbol{R}^p) \to \boldsymbol{R}$ es una función aditiva positiva que es invariante bajo la traslación de conjuntos (es decir, $\mu(x + A) = \mu(A)$ para toda $x \in \boldsymbol{R}^p$, $A \in \mathcal{D}(\boldsymbol{R}^p)$]. Demostrar que μ tiene una densidad fuerte en $\boldsymbol{R}^p$ que es una constante en $\boldsymbol{R}^p$.

(e) Si $f : \Omega \to \boldsymbol{R}$ es continua y si F está definida como en *(a)*, demostrar que F tiene densidad fuerte f en Ω.

(f) Si $G : \mathcal{D}(\Omega) \to \boldsymbol{R}$ es aditiva y tiene densidad fuerte $g : \Omega \to \boldsymbol{R}$, demostrar que g es continua en Ω. Por lo tanto, g es uniformemente continua en todo $A \in \mathcal{D}(\Omega)$.

(g) Supóngase que $G : \mathcal{D}(\Omega) \to \boldsymbol{R}$ es aditiva y tiene densidad fuerte idénticamente igual a cero en Ω. Demostrar que si K es un cubo cerrado y si $\varepsilon > 0$, entonces existe una partición de K en cubos $\{K_1, \ldots, K_r\}$ tal que $|G(K_j)| \le \varepsilon c(K_j)$ para $j = 1, \ldots, r$, por lo que se infiere que $|G(K)| \le \varepsilon c(K)$. Concluir que $G(K) = 0$ para todos los cubos cerrados $K \subseteq \Omega$.

(h) Supóngase que F_1 y F_2 son funciones aditivas en $\mathcal{D}(\Omega)$ tales que para alguna $M > 0$ se tiene $|F_j(A)| \le Mc(A)$ para toda $A \in \mathcal{D}(\Omega)$, $j = 1, 2$. Si $F_1(K) = F_2(K)$ para todo cubo $K \subseteq \Omega$, demostrar que $F_1(A) = F_2(A)$ para toda $A \in \mathcal{D}(\Omega)$.

Sección 45 Transformación de conjuntos e integrales

En la sección 43 se vio que las transformaciones continuas de un intervalo en $\boldsymbol{R}$ pueden cubrir un cubo cerrado en $\boldsymbol{R}^2$. Se probará que este fenómeno no puede suceder si la aplicación está en la clase C^1 y se estudiará la transformación de conjuntos con contenido bajo transformaciones C^1 El caso de una transformación lineal es de modo particular importante y el resultado es satisfactoriamente sencillo. En el caso de una transformación no lineal, se verá que el jacobiano de la transformación indica la extensión de la "distorsión" de la transformación.

Estos resultados se usarán para probar un teorema concerniente al "cambio de variable" de una integral sobre un conjunto en $\boldsymbol{R}^p$. Los casos especiales de coordenadas polares y esféricas se examinan brevemente y se da un teorema más fuerte aplicable a muchas transformaciones de singularidad moderada.

45.1 LEMA. *Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y suponga que $\varphi : \Omega \to \boldsymbol{R}^p$ pertenece a la clase $C^1(\Omega)$. Sea A un conjunto acotado con $A^- \subseteq \Omega$. Entonces, existe un conjunto abierto acotado Ω_1 con $A^- \subseteq \Omega_1 \subseteq \Omega_1^- \subseteq \Omega$ y una constante $M > 0$ tal que si A está contenida en la unión de un número finito de cubos cerrados en Ω_1 con contenido total a lo más α, entonces $\varphi(A)$ está contenido en la unión de un número finito de cubos cerrados con contenido total a lo más $(\sqrt{p}\,M)^p\alpha$.*

DEMOSTRACION. Si $\Omega = \boldsymbol{R}^p$, sea $\delta = 1$; si no sea $\delta = \frac{1}{2}\inf\{\|a - x\| : a \in A^-,\ x \notin \Omega\}$. Dado que A^- es compacto, se infiere que $\delta > 0$. (¿por qué?) Ahora, sea $\Omega_1 = \{y \in \boldsymbol{R}^p : \|y - a\| < \delta$ para alguna $a \in A\}$, de tal manera que Ω_1 es abierto y acotado y $A^- \subseteq \Omega_1$ y $\Omega_1^- \subseteq \Omega$. Dado que $\varphi \in C^1(\Omega)$ y Ω_1^- es compacto, se deduce que $M = \sup\{\|D\varphi(x)\|_{pp} : x \in \Omega_1\}$ es finito. Si $A \subseteq I_1 \cup \cdots \cup I_n$, en donde las I_j son cubos cerrados contenidos en Ω_1, entonces del corolario 40.6 se infiere que para $x, y \in I_j$ se tiene

$$\|\varphi(x) - \varphi(y)\| \leq M\,\|x - y\|.$$

Suponga que la longitud lateral de I_j es $2r_j$ y tómese x como el centro de I_j entonces, si $y \in I_j$, se tiene $\|x - y\| \leq \sqrt{p}\,r_j$ de donde $\|\varphi(x) - \varphi(y)\| \leq \sqrt{p}\,Mr_j$. De modo que $\varphi(I_j)$ está contenido en un cubo cerrado de longitud lateral $2\sqrt{p}\,Mr_j$. Por lo tanto, se sigue que $\varphi(A)$ está contenido en la unión de un número finito de cubos cerrados con contenido total a lo más $(\sqrt{p}\,M)^p\alpha$.

45.2 TEOREMA. *Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y supóngase que $\varphi : \Omega \to \boldsymbol{R}^p$ pertenece a la clase $C^1(\Omega)$. Si $A \subset \Omega$ tienen contenido cero y si $A^- \subseteq \Omega$, entonces $\varphi(A)$ tiene contenido cero.*

45.3 COROLARIO. *Sea* $r<p$, *sea* $\Omega\subseteq R^r$ *abierto y suponga que* $\psi:\Omega\to R^p$ *pertenece a la clase* $C^1(\Omega)$. *Si* $A\subseteq\Omega$ *es un conjunto acotado con* $A^-\subseteq\Omega$, *entonces,* $\psi(A)$ *tiene contenido cero en* R^p.

DEMOSTRACION. Sea $\Omega_0=\Omega\times R^{p-r}$ de tal manera que Ω_0 es abierto en R^p,y defínase $\varphi:\Omega_0\to R^p$ como

$$\varphi(x_1,\ldots,x_r,x_{r+1},\ldots,x_p)=\psi(x_1,\ldots,x_r).$$

Es evidente que $\varphi\in C^1(\Omega_0)$. Sea $A_0=A\times\{0,\ldots,0\}$ de tal manera que $A_0^-\subseteq\Omega_0$ y A_0 tiene contenido cero en R^p. Se sigue que $\psi(A)=\varphi(A_0)$ tiene contenido cero en R^p.

Observe que este corolario asegura que la imagen C^1 de cualquier conjunto acotado de "menor dimensionalidad" tiene contenido cero. Q.E.D.

Dado que la frontera del conjunto A con contenido tiene contenido cero, del teorema 45.2 se infiere que si φ está en la clase C^1 entonces $\varphi(b(A))$ tiene contenido cero. Desafortunadamente $\varphi(b(A))$ tiene poca relación, en general, con $b(\varphi(A))$. Esta observación incrementa el interés de los dos resultados que siguen.

45.4 TEOREMA. *Sea* $\Omega\subseteq R^p$ *abierto y suponga que* $\varphi:\Omega\to R^p$ *pertenece a la clase* $C^1(\Omega)$. *Suponga que A tiene contenido,* $A^-\subseteq\Omega$, *y* $J_\varphi(x)\neq 0$ *para toda* $x\in A^0$. *Entonces,* $\varphi(A)$ *tiene contenido.*

DEMOSTRACION. Dado que A^- es compacto y φ es continua, entonces $\varphi(A)\subseteq\varphi(A^-)$ es acotado. Para probar que $\varphi(A)$ tiene contenido se habrá de probar que $b(\varphi(A))\subseteq\varphi(b(A))$ y que $\varphi(b(A))$ tiene contenido cero.

Como $\varphi(A^-)$ es compacto, se tiene $b(\varphi(A))\subseteq\varphi(A^-)=\varphi(A^0\cup b(A))$. Por lo que, si $y\in b(\varphi(A))$,existe una $x\in A^0\cup b(A)$ tal que $y=\varphi(x)$. Si $x\in A^0$ entonces $J_\varphi(x)\neq 0$ y del teorema de aplicación suprayectiva 41.6 se infiere que $y=\varphi(x)$ es un punto interior de $\varphi(A^0)\subseteq\varphi(A)$.Pero esto contradice la hipótesis de que $y\in b(\varphi(A))$. Por lo tanto, se infiere que $b(\varphi(A))\subseteq\varphi(b(A))$.

Ahora, como A tiene contenido, su frontera $b(A)\subseteq\Omega$ es un conjunto cerrado con contenido cero; por lo tanto, del teorema 45.2 se deduce que $\varphi(b(A))$ tiene contenido cero.

45.5 COROLARIO. *Sea* $\Omega\subseteq R^p$ *abierto y supóngase que* $\varphi:\Omega\to R^p$ *pertenece a la clase* $C^1(\Omega)$ *y es inyectiva en* Ω. *Si A tiene contenido* $A^-\subseteq\Omega$, *y* $J_\varphi(x)\neq 0$ *para* $x\in A^0$, *entonces* $b(\varphi(A))=\varphi(b(A))$.

DEMOSTRACION. Basta demostrar que $\varphi(b(A))\subseteq b(\varphi(A))$, ya que la inclusión inversa se probó en la demostración del teorema. Sea $x\in b(A)$, de manera que exista una sucesión (x_n) en A y una sucesión (y_n) en

$\Omega \setminus A$, y ambas convergen a x. Dado que φ es continua $\varphi(x_n) \to \varphi(x)$ y $\varphi(y_n) \to \varphi(x)$. Como φ es inyectiva en Ω, entonces $\varphi(y_n) \notin \varphi(A)$ y $\varphi(x) \in b(\varphi(A))$. Por lo tanto $\varphi(b(A)) \subseteq b(\varphi(A))$.

Transformaciones por aplicaciones lineales

En seguida se verá que los conjuntos con contenido son aplicados por una aplicación *lineal*, en $\boldsymbol{R}^p$ hacia conjuntos cuyo contenido es un múltiplo fijo del contenido original. Más aún, este múltiplo es el valor absoluto del determinante correspondiente a la aplicación lineal. (En este teorema se habrá de suponer que al lector le son familiares el concepto y las propiedades elementales del determinante de una aplicación lineal en $\boldsymbol{R}^p$)

45.6 TEOREMA. *Sea* $L \in \mathscr{L}(\boldsymbol{R}^p)$. *Si* $A \in \mathscr{D}(\boldsymbol{R}^p)$, *entonces* $c(L(A)) = |\det L|\, c(A)$.

DEMOSTRACION. Si L es singular (es decir, si $\det L = 0$), entonces L aplica a $\boldsymbol{R}^p$ en un subespacio lineal propio de $\boldsymbol{R}^p$. Dado que este subespacio también se puede obtener como la imagen de alguna $L' : \boldsymbol{R}^r \to \boldsymbol{R}^p$ con $r < p$, del corolario 45.3 se infiere $c(L(A)) = 0$ para todo $A \in \mathscr{D}(\boldsymbol{R}^p)$. Por lo que la afirmación es válida para aplicaciones lineales que sean singulares.

Si L no es singular (es decir, si $\det L \neq 0$), entonces el teorema 45.4 implica que si $A \in \mathscr{D}(\boldsymbol{R}^p)$, entonces $L(A) \in \mathscr{D}(\boldsymbol{R}^p)$. Se define ahora $\lambda : \mathscr{D}(\boldsymbol{R}^p) \to \boldsymbol{R}$ como $\lambda(A) = c(L(A))$. *(i)* Es claro que $\lambda(A) \geq 0$ para toda $A \in \mathscr{D}(\boldsymbol{R}^p)$. *(ii)* Supóngase que $A, B \in \mathscr{D}(\boldsymbol{R}^p)$ y $A \cap B = \emptyset$; entonces

$$\lambda(A \cup B) = c(L(A \cup B)) = c(L(A) \cup L(B)).$$

Dado que L es inyectiva, $L(A) \cap L(B) = \emptyset$ y entonces

$$c(L(A) \cup L(B)) = c(L(A)) + c(L(B)) = \lambda(A) + \lambda(B).$$

(iii) Sea $x \in \boldsymbol{R}^p$ y $A \in \mathscr{D}(\boldsymbol{R}^p)$; entonces

$$\lambda(x + A) = c(L(x + A)) = c(L(x) + L(A)) = c(L(A)) = \lambda(A).$$

Por lo tanto, del corolario 44.7 se deduce que existe una constante $m_L \geq 0$ tal que $\lambda(A) = m_L c(A)$ para toda $A \in \mathscr{D}(\boldsymbol{R}^p)$. En seguida se examina cómo m_L depende de $L \in \mathscr{L}(\boldsymbol{R}^p)$. Sea $M \in \mathscr{L}(\boldsymbol{R}^p)$ no singular; entonces, si $A \in \mathscr{D}(\boldsymbol{R}^p)$, se tiene

$$\begin{aligned} m_{L \circ M} c(A) &= c(L \circ M(A)) = c(L(M(A))) \\ &= m_L c(M(A)) = m_L m_M c(A)). \end{aligned}$$

Por lo tanto, se tiene $m_{L \circ M} = m_L m_M$ para toda $L, M \in \mathscr{L}(\boldsymbol{R}^p)$ no singulares.

Queda por demostrar que $m_L = |\det L|$. Para hacer esto, del álgebra lineal se usará el hecho de que toda $L \in \mathscr{L}(\mathbf{R}^p)$ no singular es la composición de aplicaciones lineales de las siguientes tres formas:

(a) $L_1(x_1, \ldots, x_p) = (\alpha x_1, x_2, \ldots, x_p)$ para cualquiera $\alpha \neq 0$;
(b) $L_2(x_1, \ldots, x_i, x_{i+1}, \ldots, x_p) = (x_1, \ldots, x_{i+1}, x_i, \ldots, x_p)$;
(c) $L_3(x_1, \ldots, x_p) = (x_1 + x_2, x_2, \ldots, x_p)$.

Obsérvese que si K_0 es el cubo semi-abierto $[0, 1) \times \cdots \times [0, 1)$ en $\mathbf{R}^p$ y si $\alpha > 0$, entonces $L_1(K_0) = [0, \alpha) \times [0, 1) \times \cdots \times [0, 1)$, por lo que se infiere que

$$\alpha = c(L_1(K_0)) = m_{L_1} c(K_0) = m_{L_1}.$$

Análogamente, si $\alpha < 0$, entonces $L_1(K_0) = (\alpha, 0] \times [0, 1) \times \cdots \times [0, 1)$, y

$$-\alpha = c(L_1(K_0)) = m_{L_1} c(K_0) = m_{L_1}.$$

Por lo tanto, en cualquiera de los casos se tiene $m_{L_1} = |\alpha| = |\det L_1|$.

Dado que $L_2(K_0) = K_0$, se sigue que $m_{L_2} = 1 = |\det L_2|$.

Por último, sean Δ_1 y Δ_2 los dos conjuntos

$$\Delta_1 = \{(x_1, \ldots, x_p) : 0 \leq x_i < 1,\ x_1 < x_2\},$$
$$\Delta_2 = \{(x_1, \ldots, x_p) : 0 \leq x_i < 1,\ x_2 \leq x_1\}.$$

Es claro que $\Delta_1 \cap \Delta_2 = \emptyset$ y $K_0 = \Delta_1 \cup \Delta_2$. Dado que se puede ver que

$$L_3(K_0) = \Delta_2 \cup \{(1, 0, \ldots, 0) + \Delta_1\}$$

se sigue que

$$\begin{aligned} c(L_3(K_0)) &= c(\Delta_2) + c((1, 0, \ldots, 0) + \Delta_1) = c(\Delta_2) + c(\Delta_1) \\ &= c(\Delta_1 \cup \Delta_2) = c(K_0). \end{aligned}$$

Por lo tanto, $m_{L_3} = 1 = |\det L_3|$.

Ahora, sea la aplicación lineal no singular L la composición de aplicaciones lineales $L_1, L_2, \ldots, L_r$ con una de las tres formas que se dieron antes. Dado que

$$\begin{aligned} m_L &= m_{L_1 \circ L_2 \circ \cdots \circ L_r} = m_{L_1} m_{L_2} \cdots m_{L_r} \\ &= |\det L_1|\,|\det L_2| \cdots |\det L_r| \\ &= |(\det L_1)(\det L_r) \cdots (\det L_r)| \\ &= |\det (L_1 \circ L_2 \circ \cdots \circ L_r)| = |\det L|, \end{aligned}$$

el teorema está demostrado Q.E.D.

Transformación por aplicaciones no lineales

Se obtendrá ahora una extensión del teorema 45.6 para aplicaciones C^1 que no son lineales. Desde luego, en este caso el contenido de la imagen de un conjunto arbitrario no necesariamente es un múltiplo fijo del contenido del conjunto dado sino que puede variar de un punto a otro. El teorema jacobiano implica que si K es un cubo suficientemente pequeño con centro x entonces $c(\varphi(K))$ es aproximadamente igual a $|J_\varphi(x)|\,c(K)$. Este resultado es crucial para poder demostrar el teorema del cambio de variables. Será técnicamente conveniente considerar primero el siguiente caso especial.

45.7 LEMA. *Sea $K \subseteq \mathbf{R}^p$ un cubo cerrado con centro 0. Sea Ω un conjunto abierto que contiene a K y suponga que $\psi: \Omega \to \mathbf{R}^p$ pertenece a la clase $C^1(\Omega)$ y es inyectiva. Suponga además que $J_\psi(x) \neq 0$ para $x \in K$ y que*

$$\|\psi(x) - x\| \le \alpha\,\|x\| \quad \text{para } x \in K, \tag{45.1}$$

en donde α satisface $0 < \alpha < 1/\sqrt{p}$. Entonces

$$(1-\alpha\sqrt{p})^p \le \frac{c(\psi(K))}{c(K)} \le (1+\alpha\sqrt{p})^p.$$

DEMOSTRACION. Del teorema 45.4 se sigue que $\psi(K)$ tiene contenido y del corolario 45.5 se infiere que $b(\psi(K)) = \psi(b(K))$. Si la longitud lateral de K es $2r$ y si $x \in b(K)$, entonces (por el teorema 8.10) se tiene $r \le \|x\| \le r\sqrt{p}$. La desigualdad (45.1) implica que $\psi(x)$ está comprendido en una distancia $\alpha r\sqrt{p}$ de $x \in b(K)$. Por lo tanto, el conjunto compacto $\psi(b(K)) = b(\psi(K))$ no interseca a un cubo abierto C_i con centro 0 y longitud lateral $2(1-\alpha\sqrt{p})r$. Si se toma A (respectivamente B) como el conjunto de todos los puntos interiores (respectivamente, exteriores) de $\psi(K)$, entonces A y B son conjuntos abiertos no vacíos con unión $\mathbf{R}^p \setminus b(\psi(K))$. Dado que C_i es conexo en $\mathbf{R}^p$, se debe tener $C_i \subseteq A$ o bien $C_i \subseteq B$. Pero dado que $0 \in C_i \cap A$, se infiere que $C_i \subseteq A \subseteq \psi(K)$. De manera análoga, el lector puede probar que si C_o es el cubo cerrado con centro 0 y longitud lateral $2(1+\alpha\sqrt{p})r$, entonces $\psi(K) \subseteq C_o$. La conclusión establecida se deduce ahora de estas inclusiones.

45.8 EL TEOREMA JACOBIANO. *Sea $\Omega \subseteq \mathbf{R}^p$ abierto y suponga que $\varphi: \Omega \to \mathbf{R}^p$ pertenece a la clase $C^1(\Omega)$, es inyectiva en Ω, y que $J_\varphi(x) \neq 0$ para $x \in \Omega$. Suponga que A tiene contenido y $A^- \subseteq \Omega$. Si $\varepsilon > 0$ está dada, entonces existe $\gamma > 0$ tal que si K es un cubo cerrado con centro $x \in A$ y longitud lateral menor a 2γ, entonces*

$$|J_\varphi(x)|\,(1-\varepsilon)^p \le \frac{c(\varphi(K))}{c(K)} \le |J_\varphi(x)|\,(1+\varepsilon)^p. \tag{45.2}$$

DEMOSTRACION. Construir $\delta > 0$ y Ω_1 como en la demostración del

lema 45.1 Dado que det $D\varphi(x) = J_\varphi(x) \neq 0$ para toda $x \in \Omega$, se infiere que existe $L_x = (D\varphi(x))^{-1}$ puesto que $1 = \det\,(L_x \circ D\varphi(x)) = (\det L_x)(\det D\varphi(x))$, se deduce que

$$\det L_x = 1/J_\varphi(x) \qquad \text{para } x \in \Omega.$$

Dado que en la representación matricial típica para L_x las entradas son funciones continuas, de la compacidad de Ω_1^- y (21.4) se infiere que existe una constante $M > 0$ tal que $\|L_x\|_{pp} \leq M$ para toda $x \in \Omega_1$.

Sea ε, dada con $0 < \varepsilon < 1$, Dado que la aplicación $x \mapsto D\varphi(x)$ es uniformemente continua en Ω_1, existe β con $0 < \beta < \delta$ tal que si $x_1, x_2 \in \Omega_1$ y $\|x_1 - x_2\| \leq \beta$; entonces, $\|D\varphi(x_1) - D\varphi(x_2)\|_{pp} \leq \varepsilon/M\sqrt{p}$. Sea ahora $x \in A$ dada, de donde, si $\|z\| \leq \beta$, entonces x y $x + z$ pertenecen a Ω_1. Por lo tanto, del lema 41.3 se infiere que

$$(45.3) \quad \|\varphi(x+z) - \varphi(x) - D\varphi(x)(z)\| \leq \|z\| \sup_{0 \leq t \leq 1} \|D\varphi(x+tz) - D\varphi(x)\|_{pp}$$

$$\leq \frac{\varepsilon}{M\sqrt{p}} \|z\|.$$

Sea $x \in A$ y defínase $\psi(z)$ para $\|z\| \leq \beta$ como

$$\psi(z) = L_x[\varphi(x+z) - \varphi(x)].$$

Dado que $L_x = (D\varphi(x))^{-1}$, la desigualdad (45.3) da

$$\|\psi(z) - z\| \leq \frac{\varepsilon}{\sqrt{p}} \|z\| \qquad \text{para } \|z\| \leq \beta.$$

Se aplica ahora el lema anterior con $\alpha = \varepsilon/\sqrt{p}$ para inferir que si K_1 es cualquier cubo cerrado con centro 0 y contenido en una bola abierta con radio β, entonces

$$(1-\varepsilon)^p \leq \frac{c(\psi(K_1))}{c(K_1)} \leq (1+\varepsilon)^p.$$

De la definición de ψ y del teorema 45.6 se infiere que si $K = x + K_1$, entonces K es un cubo cerrado con centro x y que $c(K) = c(K_1)$ y

$$c(\psi(K_1)) = |\det L_x|\, c(\varphi(x + K_1) - \varphi(x))$$

$$= \frac{1}{|J_\varphi(x)|} c(\varphi(K)).$$

Por lo tanto, si K es un cubo cerrado con centro $x \in A$ y longitud lateral menor a 2γ (en donde $\gamma = \beta/\sqrt{p}$), la desigualdad (45.2) es válida Q.E.D.

Cambio de variables

Se aplicará ahora el teorema jacobiano para obtener un importante teorema que es una generalización a $\mathbf{R}^p$ del teorema de cambio de variables

30.12. El último resultado segura que si $\varphi:[\alpha, \beta] \to \mathbf{R}$ tiene una derivada continua y si f es continua en el rango de φ, entonces

$$\int_{\varphi(\alpha)}^{\varphi(\beta)} f = \int_{\alpha}^{\beta} (f \circ \varphi)\varphi'. \tag{45.4}$$

El resultado que se va a probar trata a una aplicación inyectiva φ definida en un subconjunto abierto $\Omega \subseteq \mathbf{R}^p$ con valores en $\mathbf{R}^p$. Se habrá de suponer que $\varphi \in C^1(\Omega)$ y que su determinante jacobiano

$$J_\varphi(x) = \det [D_j\varphi_i(x)]$$

no es cero en Ω. Se probará que si A tiene contenido, si $A^- \subseteq \Omega$, y si f es acotada y continua de $\varphi(A)$ a $\mathbf{R}$, entonces $\varphi(A)$ tiene contenido y

$$\int_{\varphi(A)} f = \int_A (f \circ \varphi)\,|J_\varphi|. \tag{45.5}$$

Se podrá ver que en cierto modo, las hipótesis son más restrictivas que en el caso $p = 1$. De hecho, en (45.4) no se supone que φ sea inyectiva o que $\varphi'(x) \neq 0$ para $x \in [\alpha, \beta]$. Obsérvese que si φ fuera inyectiva, el ángulo exacto para (45.5) en el caso $p = 1$ sería

$$\int_A^B f = \int_\alpha^\beta (f \circ \varphi)\,|\varphi'|,$$

en donde $A = \inf\{\varphi(\alpha), \varphi(\beta)\}$ y $B = \sup\{\varphi(\alpha), \varphi(\beta)\}$. Desde luego, si $\varphi'(x) > 0$ para $\alpha \le x \le \beta$, entonces, la fórmula (45.5) se reduce a (45.4), mientras que si $\varphi'(x) < 0$ para $\alpha \le x \le \beta$, la fórmula (45.5) se reduce a

$$\int_{\varphi(\beta)}^{\varphi(\alpha)} f = \int_\alpha^\beta (f \circ \varphi)(-\varphi'),$$

por lo que (45.4) también se infiere. La aplicación de esta diferencia es que la integral sobre intervalos en $\mathbf{R}$ está "orientada" en el sentido que se define

$$\int_u^v f = -\int_v^u f$$

para cualesquiera números reales u, v. Ninguna orientación de este tipo se ha definido para integrales sobre $\mathbf{R}^p$.

La demostración que aquí se da se debe esencialmente a J. T. Schwartz†. Es "elemental" en el sentido de que no se hace uso de ninguno de los resultados de teoría de la medida. Sin embargo, el argumento es muy deli-

† J. T. SCHWARTZ (1930-) se graduó de CCNY, recibió su doctorado en la Universidad de Yale y es profesor en el Instituto Courant de la Universidad de Nueva York. Aun cuando se le conoce básicamente por su trabajo en análisis funcional, también ha hecho aportaciones en ecuaciones diferenciales, geometría, lenguajes de computación, varios aspectos de física matemática y en economía matemática.

cado y se utilizan varias propiedades más profundas de funciones continuas, conjuntos compactos y conexos y las propiedades de la integral. Aun así, el teorema que se demostrará no es del todo suficiente para todos los casos importantes que surgen de tal manera que se discutirá en seguida con un argumento más fuerte que hace posible que J_φ sea cero y $f \circ \varphi$ sea discontinua en un conjunto con contenido cero.

45.9 TEOREMA DEL CAMBIO DE VARIABLE. *Sea* $\Omega \subseteq \mathbf{R}^p$ *abierto y supóngase que* $\varphi : \Omega \to \mathbf{R}^p$ *pertenece a la clase* $C^1(\Omega)$, *es inyectiva en* Ω, *y* $J_\varphi(x) \neq 0$ *para* $x \in \Omega$. *Supóngase que* A *tiene contenido,* $A^- \subseteq \Omega$, *y* $f : \varphi(A) \to \mathbf{R}$ *es acotada y continua. Entonces*

$$\int_{\varphi(A)} f = \int_A (f \circ \varphi)\,|J_\varphi|. \tag{45.5}$$

DEMOSTRACION. Del teorema 45.4 se infiere que $\varphi(A)$ tiene contenido. Dado que los integrandos son continuos, se deduce que existen las integrales en (45.5); queda por demostrar su igualdad. Si se toma $f = f^+ - f^-$, en donde $f^+ = \frac{1}{2}(f + |f|)$ y $f^- = \frac{1}{2}(|f| - f)$, se usa la linealidad de la integral, basta con suponer que $f(y) \geq 0$ para toda $y \in \varphi(A)$.

Tómese Ω_1 como en el lema 45.1 y sea

$$M_\varphi = \sup\{\|D\varphi(x)\|_{pp} : x \in \Omega_1\},$$
$$M_f = \sup\{f(y) : y \in \varphi(A)\},$$
$$M_J = \sup\{|J_\varphi(x)| : x \in A\}.$$

Sea $\varepsilon > 0$ arbitraria, excepto que $0 < \varepsilon < 1$, sea I una celda cerrada que contiene a A y sea $\{K_i : i = 1, \ldots, M\}$ una partición de I en cubos cerrados no traslapados con longitud lateral menor a 2γ, en donde γ es la constante del teorema jacobiano. Supóngase que los cubos contenidos completamente en A se enumeran $K_1, \ldots, K_m$; aquellos que tienen puntos tanto en A como en su complemento se enumeran $K_{m+1}, \ldots, K_n$, y los cubos completamente contenidos en el complemento de A se enumeran $K_{n+1}, \ldots, K_M$. Dado que A tiene contenido, se puede suponer que la partición se ha elegido lo suficientemente fina como para que

$$c(A) \leq \sum_{i=1}^{m} c(K_i) + \varepsilon, \qquad \sum_{i=m+1}^{n} c(K_i) < \varepsilon. \tag{i}$$

Se toma $B = K_1 \cup \cdots \cup K_m$ de tal manera que $B \subseteq A$. Dado que $c(A \setminus B) = c(A) - c(B) <$ se tiene ε,

$$\left| \int_A (f \circ \varphi)\,|J_\varphi| - \int_B (f \circ \varphi)\,|J_\varphi| \right| = \left| \int_{A \setminus B} (f \circ \varphi)\,|J_\varphi| \right| \leq M_f M_J c(A \setminus B) \leq [M_f M_J]\varepsilon. \tag{ii}$$

Del lema 45.1 se infiere que $c(\varphi(A \setminus B)) \leq (\sqrt{p}\, M_\varphi)^p \varepsilon$, de tal manera que

(iii) $$\left| \int_{\varphi(A)} f - \int_{\varphi(B)} f \right| = \left| \int_{\varphi(A \setminus B)} f \right| \leq [M_f (\sqrt{p}\, M_\varphi)^p] \varepsilon.$$

Si x_i es el centro de K_i, $i = 1, \ldots, m$, entonces del teorema jacobiano de deduce que

$$|J_\varphi(x_i)|\,(1-\varepsilon)^p \leq \frac{c(\varphi(K_i))}{c(K_i)} \leq |J_\varphi(x_i)|\,(1+\varepsilon)^p.$$

Ahora, como $0 < \varepsilon < 1$, se puede ver que $1 - 2^p\varepsilon \leq (1-\varepsilon)^p$ y $(1+\varepsilon)^p \leq 1 + 2^p\varepsilon$, y esta desigualdad se puede escribir en la forma

(iv) $$|c(\varphi(K_i)) - |J_\varphi(x_i)|\, c(K_i)| \leq [c(K_i) M_J 2^p] \varepsilon.$$

Dada la continuidad de las funciones en el integrando en el conjunto compacto B, se infiere que se puede suponer que para cualquier punto $y_i \in K_i$,

(v) $$\left| \int_B (f \circ \varphi)\, |J_\varphi| - \sum_{i=1}^{m} (f \circ \varphi)(y_i)\, |J_\varphi(x_i)|\, c(K_i) \right| < \varepsilon c(B).$$

(Ya que, si fuera necesario, se podrán dividir los cubos $K_1, \ldots, K_m$ en cubos pequeños. Ver el ejercicio 43.T.)

Dado que φ es inyectiva, dos conjuntos de $\{\varphi(K_i) : i = 1, \ldots, m\}$ se intersecan cuando más en un conjunto $\varphi(K_i \cap K_j)$ que tiene contenido 0 ya que $c(K_i \cap K_j) = 0$. Además, dado que $\varphi(K_i)$ tiene contenido, f es integrable en $\varphi(K_i)$; por lo tanto, del teorema 44.9(b) se infiere que

$$\int_{\varphi(B)} f = \sum_{i=1}^{m} \int_{\varphi(K_i)} f.$$

Ahora, como K_i es conexo, $\varphi(K_i)$ es conexo. Dado que f es acotada y continua en $\varphi(K_i)$, del teorema del valor medio 44.11 se deduce que existe $p_i \in \varphi(K_i)$ tal que

$$\int_{\varphi(K_i)} f = f(p_i) c(\varphi(K_i)), \qquad i = 1, \ldots, m.$$

Dado que $p_i \in \varphi(K_i)$, existe una única $y_i \in K_i$ con $p_i = \varphi(y_i)$, $i = 1, \ldots, m$. Por lo tanto, se tiene

(vi) $$\int_{\varphi(B)} f = \sum_{i=1}^{m} (f \circ \varphi)(y_i) c(\varphi(K_i)).$$

Pero puesto que $(f \circ \varphi)(y_i) \geq 0$, de *(iv)* se infiere que

$$\left| \sum_{i=1}^{m} (f \circ \varphi)(y_i) c(\varphi(K_i)) - \sum_{i=1}^{m} (f \circ \varphi)(y_i)\, |J_\varphi(x_i)|\, c(K_i) \right|$$
$$\leq \left[M_J 2^p \sum_{i=1}^{m} (f \circ \varphi)(y_i) c(K_i) \right] \varepsilon$$

$$\leq \left[M_J M_f 2^p \sum_{i=1}^{m} c(K_i)\right]\varepsilon \leq [M_J M_f 2^p c(A)]\varepsilon.$$

Si se combina esta última relación con *(v)* y *(vi)* se obtiene

(vii) $$\left| \int_{\varphi(B)} f - \int_B (f \circ \varphi)\, |J_\varphi| \right| \leq (1 + M_J M_f 2^p) c(A) \varepsilon.$$

Combinando *(vii)* con *(ii)* y *(iii)*, se obtiene

$$\left| \int_{\varphi(A)} f - \int_A (f \circ \varphi)\, |J_\varphi| \right| \leq [M_f(\sqrt{p} M_\varphi)^p + (1 + M_J M_f 2^p) c(A) + M_f M_J]\varepsilon.$$

Dado que ε es un número arbitrario con $0 < \varepsilon < 1$, la ecuación (45.5) queda probada

Q.E.D.

Aplicaciones

El *uso* del teorema sobre el cambio de variables cuando $p > 1$ por lo general es diferente de la aplicación del teorema correspondiente cuando $p = 1$. Por ejemplo, al calcular

$$\int_0^1 x(1 + x^2)^{1/2}\, dx$$

por lo común se nota que si se introduce $\varphi(x) = 1 + x^2$ entonces $\varphi'(x) = 2x$; por lo que el integrando tiene la forma $\frac{1}{2}(\varphi(x))^{1/2}\varphi'(x)$ y entonces

$$\int_0^1 x(1 + x^2)^{1/2}\, dx = \tfrac{1}{2}\tfrac{2}{3}(\varphi(x))^{3/2}\Big|_{x=0}^{x=1}$$

$$= \tfrac{1}{3}(1 + x^2)^{3/2}\Big|_{x=0}^{x=1} = \tfrac{1}{3}(2^{3/2} - 1).$$

De modo que la integración se lleva a cabo observando que el integrando dado es una composición de alguna función y φ, multiplicada por la derivada de φ. Aplicaciones análogas para calcular integrales en más de una variable por lo general son posibles sólo cuando el término jacobiano es constante (o muy sencillo). Por ejemplo, una integral de la forma

$$\iint_A f(x + 2y, 2x - 3y)\, d(x, y)$$

se puede tratar introduciendo la transformación lineal $\varphi(x, y) = (x + 2y, 2x - 3y)$. En este caso

$$J_\varphi(x, y) = \det \begin{bmatrix} 1 & 2 \\ 2 & -3 \end{bmatrix} = -3 - 4 = -7$$

y entonces se tiene

$$\iint_A f(x+2y, 2x-3y)\, d(x, y) = \frac{1}{7} \iint_{\varphi(A)} f(u, v)\, d(u, v).$$

Esta segunda integral puede ser más sencilla si $f(u, v)$ es más sencilla (por ejemplo, si $f(u, v) = g(u)h(v)$], o si $\varphi(A)$ es más sencilla (por ejemplo, si es una celda). De otra manera, puede no simplificar mucho las cosas.

Un uso más típico del teorema es para calcular una integral múltiple $\int_D f$ observando que el conjunto D es la imagen de un conjunto más sencillo A (por ejemplo, una celda) con una aplicación apropiada φ.

45.10 EJEMPLOS. *(a)* Supóngase que D designa al rectángulo con vértices (0,0), (2,2), (1,3), (−1, 1); es decir, la región acotada por las rectas dadas por

$$y = x, \quad y = -x+4, \quad y = x+2, \quad y = -x.$$

Si se toma $u = y - x$ y $v = y + x$, estas rectas se transforman en

$$u = 0, \quad v = 4, \quad u = 2, \quad v = 0.$$

Por lo tanto, si φ es la aplicación $\varphi(u, v) = (x, y)$, entonces φ aplica la celda $A = [0, 2] \times [0, 4]$ en D. Se le deja al lector demostrar que

$$\iint_D f(x, y)\, d(x, y) = \iint_A f[\tfrac{1}{2}(v-u), \tfrac{1}{2}(u+v)]\tfrac{1}{2}\, d(u, v)$$

$$= \frac{1}{2}\int_0^4 \left\{ \int_0^2 f[\tfrac{1}{2}(v-u), \tfrac{1}{2}(u+v)]\, du \right\} dv.$$

(b) Sea $D \subseteq \boldsymbol{R}^2$ el conjunto de puntos en $\boldsymbol{R}^2$ dado por

$$D = \{(u, v) : 1 \le u^2 - v^2 \le 9,\ 1 \le uv \le 4\};$$

De donde, D está acotado por cuatro hipérbolas. Si se define $\psi : (u, v) \mapsto (x, y)$ como

$$x = u^2 - v^2, \qquad y = uv,$$

entonces es claro que ψ aplica estas hipérbolas del plano (u, v) en las rectas $x = 1$, $x = 9$, $y = 1$, $y = 4$ del plano (x, y). A pesar de que ψ no es inyectiva en todo $\boldsymbol{R}^2$, es inyectiva en el conjunto $Q = \{(u, v) : u > 0, v > 0\}$ y $J_\psi(u, v) = 2(u^2 + v^2)$. Más aún, $\psi(Q) = \{(x, y) : x \in \boldsymbol{R}, y > 0\}$.

Por lo que se define φ de $\{(x, y) : x \in \boldsymbol{R}, y > 0\}$ a $Q \subseteq \boldsymbol{R}^2$ como el inverso de ψ. De lo anterior es claro que φ aplica las rectas $x = 1$, $x = 9$, $y = 1$, $y = 4$ en las hipérbolas

$$u^2 - v^2 = 1, \quad u^2 - v^2 = 9, \quad uv = 1, \quad uv = 4,$$

respectivamente, y que el conjunto D es la imagen bajo φ de la celda $A = [1, 9] \times [1, 4]$. Mediante cálculos directos se prueba que φ es de la forma $\varphi(x, y) = (u, v)$ en donde

$$(45.6) \qquad u = \left[\frac{x + (x^2 + 4y^2)^{1/2}}{2}\right]^{1/2}, \qquad v = \left[\frac{-x + (x^2 + 4y^2)^{1/2}}{2}\right]^{1/2}.$$

De aquí se infiere que $u^2 + v^2 = (x^2 + 4y^2)^{1/2}$ de tal manera que $J_\varphi(x, y) = \frac{1}{2}(x^2 + 4y^2)^{-1/2}$. Este hecho también se deduce de la identidad $(u^2 + v^2)^2 = (u^2 - v^2)^2 + 4u^2v^2 = x^2 + 4y^2$.] De modo que se tiene

$$\iint_D f(u, v)\, d(u, v) = \iint_A \frac{f(u(x, y), v(x, y))}{2\sqrt{x^2 + 4y^2}}\, d(x, y),$$

en donde $A = [1, 9] \times [1, 4]$ y en donde $u(x, y)$ y $v(x, y)$ están dadas en (45.6).

Coordenadas polares y esféricas

A menudo es conveniente especificar puntos en el plano $\mathbf{R}^2$ dando sus "coordenadas polares". Por lo general se piensa que el plano posee las coordenadas cartesianas (dadas por rectas verticales y horizontales), así como el sistema polar (dado por rayos que pasan por el origen y círculos con centro en el origen). Alternativamente, se puede pensar en las coordenadas polares como una aplicación de $(r, \theta) \in \mathbf{R}^2$ en $(x, y) \in \mathbf{R}^2$ dada por

$$(45.7) \qquad (x, y) = \varphi(r, \theta) = (r \cos \theta, r \operatorname{sen} \theta).$$

A cualquier par de números $(r, \theta) \in \mathbf{R}^2$ tales que $(x, y) = (r \cos \theta, r \operatorname{sen} \theta)$ se le llama un conjunto de **coordenadas polares** del punto (x, y). Por lo general se requiere que $r \geq 0$; aún así, cada punto (x, y) en $\mathbf{R}^2$ tiene una infinidad de conjuntos de coordenadas polares.

Por ejemplo, si $(x, y) = (0, 0)$, $(0, \theta)$ es un conjunto de coordenadas polares de $(0,0)$ para toda $\theta \in \mathbf{R}$; si $(x, y) \neq (0, 0)$ y (r, θ) es un conjunto de coordenadas polares para (x, y), entonces para cada $n \in \mathbf{Z}$ el par $(r, \theta + n2\pi)$ también es un conjunto de coordenadas polares para (x, y)

Si $(x, y) \neq (0, 0)$, entonces el par único (r, θ) con $r > 0$, $0 \leq \theta < 2\pi$, se llama el **conjunto principal de coordenadas polares del punto** (x, y). De modo que la función φ da origen a una aplicación inyectiva de $(0, +\infty) \times [0, 2\pi)$ sobre $\mathbf{R}^2 \setminus \{(0, 0)\}$. También da una aplicación de $[0, +\infty) \times [0, 2\pi)$ sobre $\mathbf{R}^2$ pero no es inyectiva, ya que manda a todos los puntos $(0, \theta)$, $0 \leq \theta < 2\pi$ en $(0,0)$. Obsérvese también que el jacobiano está dado por

$$(45.8) \qquad J_\varphi(r, \theta) = \det \begin{bmatrix} \cos \theta & -r \operatorname{sen} \theta \\ \operatorname{sen} \theta & r \cos \theta \end{bmatrix} = r(\cos \theta)^2 + r(\operatorname{sen} \theta)^2 = r,$$

que es cero para $r = 0$.

Es claro que φ aplica la celda $A = [0, 1] \times [0, 2\pi]$ del plano (r, θ) en el disco unitario $D = \{(x, y) : x^2 + y^2 \leq 1\}$ pero puesto que φ no es inyectiva en A y ya que J_φ es cero para $r = 0$, no se puede aplicar el teorema de cambio de variables 45.9 para convertir integración sobre D en integración sobre A.

Dificultades análogas se encuentran con coordenadas esféricas en $\boldsymbol{R}^3$ Recuérdese que las coordenadas esféricas están definidas por la aplicación $\Phi: \boldsymbol{R}^3 \to \boldsymbol{R}^3$ en donde

$$\text{(45.9)} \qquad \Phi(r, \theta, \phi) = (r \cos\theta \operatorname{sen}\phi,\ r \operatorname{sen}\theta \operatorname{sen}\phi,\ r\cos\phi).$$

Cualquier terna de números $(r, \theta, \phi) \in \boldsymbol{R}^3$ tales que $(x, y, z) = \Phi(r, \theta, \phi)$ se llama un conjunto de **coordenadas esféricas** de (x, y, z). Por lo general se requiere que $r \geq 0$, pero aun con esta restricción cada punto en $\boldsymbol{R}^3$ tiene una infinidad de conjuntos de coordenadas esféricas.

Por ejemplo, si $(x, y, z) = (0, 0, 0)$, entonces $(0, \theta, \phi)$ es un conjunto de coordenadas esféricas para toda $\theta \in \boldsymbol{R}$, $\phi \in \boldsymbol{R}$; si $(x, y, z) \neq (0, 0, 0)$ y (r, θ, ϕ) es un conjunto de coordenadas polares para (x, y, z), entonces para cada m, $n \in \boldsymbol{Z}$, las ternas $(r, \theta + 2m\pi, \phi + 2n\pi)$ y $(r, \theta + (2m+1)\pi, \phi + (2n+1)\pi)$ son conjuntos de coordenadas esféricas para este punto.

Si (x, y, z) es tal que $(x, y) \neq (0, 0)$, entonces la terna única (r, θ, ϕ) con $r > 0$, $0 \leq \theta < 2\pi$, $0 < \phi < \pi$, se llama **conjunto principal de coordenadas esféricas** de (x, y, z). De modo que la función Φ da una aplicación inyectiva de $(0, +\infty) \times [0, 2\pi) \times (0, \pi)$ sobre $\boldsymbol{R}^3 \setminus \{(0, 0, z) : z \in \boldsymbol{R}\}$. La restricción de Φ a $[0, +\infty) \times [0, 2\pi] \times [0, \pi]$ da una aplicación sobre todo $\boldsymbol{R}^3$ pero no es inyeciva, ya que manda a todos los puntos $(0, \theta, \phi)$ en $(0,0,0)$ y si $\phi = 0$ o π, entonces todos los puntos (r, θ, ϕ) se aplican en $(0, 0, r\cos\phi)$. Obsérvese también que

$$\text{(45.10)} \qquad J_\Phi(r, \theta, \phi) = \det \begin{bmatrix} \cos\theta \operatorname{sen}\phi & -r \operatorname{sen}\theta \operatorname{sen}\phi & r\cos\theta\cos\phi \\ \operatorname{sen}\theta \operatorname{sen}\phi & r\cos\theta \operatorname{sen}\phi & r \operatorname{sen}\theta \cos\phi \\ \cos\phi & 0 & -r \operatorname{sen}\phi \end{bmatrix}$$
$$= -r^2 \operatorname{sen}\phi.$$

Es fácil ver que Φ aplica la celda $A = [0, 1] \times [0, 2\pi] \times [0, \pi]$ del espacio (r, θ, ϕ) en la bola unitaria $D = \{(x, y, z) : x^2 + y^2 + z^2 \leq 1\}$, pero dado que Φ no es inyectiva en A y J_Φ es cero cuando r^2 sen $\phi = 0$, no se puede usar en el teorema de cambio de variables de 25.9 para convertir la integración sobre D en integración sobre A.

Se ofrecerá ahora un teorema que hace posible manejar las dificultades que se han encontrado en el uso de coordenadas polares y esféricas y que a menudo es útil en otras "transformaciones con singularidades". Se podrá ver que en el teorema no se requiere que φ sea inyectiva en el conjunto A, aunque sí es inyectiva en A^0.

45.11 TEOREMA DEL CAMBIO DE VARIABLES (FORMA FUERTE). *Sea* $\Omega \subseteq \mathbf{R}^p$ *abierto y supóngase que* $\varphi : \Omega \to \mathbf{R}^p$ *pertenece a la clase* $C^1(\Omega)$. *Sea* Ω_0 *un conjunto abierto con contenido tal que* $\Omega_0^- \subseteq \Omega$ *y tal que* φ *es inyectiva en* Ω_0. *Sea* $E \subset \Omega$ *un conjunto compacto con contenido cero tal que* $J_\varphi(x) \neq 0$ *para* $x \in \Omega_0 \setminus E$. *Supóngase que* $A \subseteq \Omega$ *tiene contenido,* $A^- \subseteq \Omega_0^-$, *que* $f : \varphi(A) \to \mathbf{R}$ *es acotada y que* f *es continua en* $\varphi(A \setminus E)$. *Entonces,*

$$\int_{\varphi(A)} f = \int_A (f \circ \varphi)\,|J_\varphi|. \tag{45.5}$$

DEMOSTRACION. Dado que $b(A)$ y $b(\Omega_0)$ son compactos y tienen contenido cero, se puede suponer que están contenidos en E; por lo tanto, $A^0 \setminus E \subseteq \Omega_0 \setminus E$. Dado que A y E tienen contenido, el conjunto $A \setminus E$ tiene contenido; más aún, como E es cerrado, $(A \setminus E)^0 = A^0 \setminus E$ de tal manera que $J_\varphi(x) \neq 0$ para $x \in (A \setminus E)^0$. Por lo tanto, por el teorema 45.4 aplicado a $A \setminus E$, se deduce que $\varphi(A \setminus E)$ tiene contenido. Del teorema 45.2 se infiere que $\varphi(E)$ tiene contenido cero y como $\varphi(A) = \varphi((A \setminus E) \cup (A \cap E)) = \varphi(A \setminus E) \cup \varphi(A \cap E)$, que $\varphi(A)$ tiene contenido. Dado que f es acotada en $\varphi(A)$ y continua excepto en un subconjunto de $\varphi(E)$, se deduce que f es integrable sobre $\varphi(A)$. Más aún, como $f \circ \varphi$ es continua excepto en un subconjunto de E, se deduce que $(f \circ \varphi)\,|J_\varphi|$ es integrable sobre A. Queda por demostrar que estas integrales son iguales.

Ahora, aplicar el lema 45.1 a E para obtener un conjunto acotado abierto Ω_1 con $E \subseteq \Omega_1 \subseteq \Omega_1^- \subseteq \Omega$ y una constante $M_1 > 0$, con la propiedad de que si E está contenido en una unión finita de cubos cerrados en Ω_1 con contenido a lo más $\alpha > 0$, entonces $\varphi(E)$ está contenido en una unión finita correspondiente de cubos cerrados con contenido a lo más $(\sqrt{p}\, M_1)^p \alpha$.

Sea $\varepsilon > 0$ dada y enciérrese E en una unión finita U_ε de cubos abiertos en Ω_1 con $c(U_\varepsilon) < \varepsilon$ y tal que la unión W_ε de las cerraduras de los cubos en U_ε siga estando contenida en Ω_1. Entonces, $c(W_\varepsilon) < \varepsilon$ y del lema 45.1 se infiere que $c(\varphi(U_\varepsilon)) \leq c(\varphi(W_\varepsilon)) \leq (\sqrt{p}\, M_1)^p \varepsilon$. Se toma ahora $B = A \setminus U_\varepsilon$ de tal manera que B tiene contenido. Ahora, dado que U_ε es abierto y contiene a $b(\Omega_0)$ y E, se infiere que $B^- \subseteq \Omega_0 \setminus E$. Se aplica ahora el teorema del cambio de variables 45.9 a B, $\Omega_0 \setminus E$ en el lugar de A, Ω, para obtener

$$\int_{\varphi(B)} f = \int_B (f \circ \varphi)\,|J_\varphi|.$$

Con facilidad se puede ver que $\varphi(A) \setminus \varphi(B) \subseteq \varphi(A \cap U_\varepsilon)$ por lo que

$$\left| \int_{\varphi(A)} f - \int_{\varphi(B)} f \right| \leq \left| \int_{\varphi(A \cap U_\varepsilon)} f \right| \leq M_f c(\varphi(A \cap U_\varepsilon))$$

$$\leq (\sqrt{p}\, M_1)^p M_f \varepsilon.$$

De manera análoga se tiene

$$\left| \int_A (f\circ\varphi)\,|J_\varphi| - \int_B (f\circ\varphi)\,|J_\varphi| \right| \le \int_{A\cap U_\varepsilon} (f\circ\varphi)\,|J_\varphi|$$
$$\le M_f M_J c(A\cap U_\varepsilon) \le M_f M_J \varepsilon.$$

Se infiere que

$$\left| \int_{\varphi(A)} f - \int_A (f\circ\varphi)\,|J_\varphi| \right| \le [(\sqrt{p}\,M_1)^p M_f + M_f M_J]\varepsilon.$$

Dado que $\varepsilon > 0$ es arbitraria, se deduce la conclusión.

Para las coordenadas polares se toma Ω_0 como un conjunto abierto con contenido en $(0, +\infty)\times(0, 2\pi)$. Para las coordenadas esféricas se toma Ω_0 como un conjunto abierto con contenido en $(0, +\infty)\times(0, 2\pi)\times(0, \pi)$.

Ejercicios

45.A. Sea $\Omega \subseteq \boldsymbol{R}^p$ un conjunto abierto y supongase que $f:\Omega\to\boldsymbol{R}^q$ satisface la condición de Lipschitz en Ω; es decir, para alguna $M>0, \|f(x)-f(y)\| \le M\,\|x-y\|$ para todas $x, y\in\Omega$. Si $K\subseteq\Omega$ es un cubo con longitud lateral $s>0$, demostrar que $f(K)$ está contenido en un cubo con longitud lateral $M\sqrt{p}s$. Demostrar que si $A\subseteq\Omega$ es un conjunto compacto con contenido cero, entonces $f(A)$ tiene contenido cero y si $B\subseteq\Omega$ es un conjunto compacto con contenido entonces $f(B)$ tiene contenido.

45.B. Considérese la aplicación de coordenadas polares $(x, y) = \varphi(r, \theta) = (r\cos\theta, r\operatorname{sen}\theta)$ sen definida en $\boldsymbol{R}^2$, y su comportamiento en el conjunto $A=[0,1]\times[0,2\pi]$. Usar el teorema 45.4 para obtener la información que asegura que la imagen $D=\varphi(A)$, que es el disco unitario $D=\{(x, y): x^2+y^2\le 1\}$, tiene contenido. Investigar la forma en que φ aplica la frontera de A. Demostrar que la frontera de D es la imagen bajo φ de sólo un lado de A y que los otros tres lados de A son aplicados al interior de D.

45.C. Considérese la aplicación de $(x, y) = \psi(u, v) = (\operatorname{sen} u, \operatorname{sen} v)$ definida en $\boldsymbol{R}^2$. Determinar la imagen de la frontera de $B=[-\frac{3}{4}\pi, \frac{3}{4}\pi]\times[-\frac{3}{4}\pi, \frac{3}{4}\pi]$ bajo ψ, y la frontera de $\psi(B)$. Demostrar que la mayoría mas no todos los puntos frontera de $\psi(B)$ son imágenes de puntos interiores de B.

45.D. Dado que el área del disco circular $\{(x, y): x^2+y^2\le 1\}$ es igual a π, encontrar las áreas de los discos elípticos dados por:

(a) $\left\{(x, y): \dfrac{x^2}{4}+\dfrac{y^2}{9}\le 1\right\}$;

(b) $\{(x, y): 2x^2+2xy+5y^2\le 1\}$.

(Sugerencia: $2x^2+2xy+5y^2=(x+2y)^2+(x-y)^2$.)

45.E. Sea B el conjunto $\{(x, y): 0\le x, 0\le y, 1\le x+y\le 2\}$. Sean $u=x+y$, $v=y$ de tal manera que B es la imagen bajo la aplicación $(x, y)=\varphi(u, v)=(u-v, v)$ del trapezoide $C=\{(u, v): 1\le u\le 2, 0\le v\le u\}$.
Demostrar que φ es inyectiva en todo $\boldsymbol{R}^2$ y que $J_\varphi(u, v)=1$. Deducir que

$$\iint_B (x+y)\,d(x,y) = \iint_C u\,d(u,v) = \tfrac{7}{3}.$$

45.F. Sea $B = \{(u,v): 0 \le u+v \le 2, 0 \le v-u \le 2\}$. Usando la transformación $(x,y) \mapsto (u,v) = (x-y, x+y)$, calcular la integral

$$\iint_B (v^2-u^2)e^{(u^2+v^2)/2}\,d(u,v).$$

45.G. Calcular la integral iterada

$$\int_1^3 \left\{ \int_{x^2}^{x^2+1} xy\,dy \right\} dx$$

directamente. Después usar la transformación $(x,y) \mapsto (u,v) = (x, y-x^2)$ para calcular esta integral.

45.H. Determinar el área de la región acotada por las curvas

$$xy = 1, \quad xy = 2, \quad y = x^2, \quad y = 2x^2$$

introduciendo un cambio de variable apropiado.

45.I. Supóngase que $\psi: \mathbf{R}^2 \to \mathbf{R}^2$ está definida como $(u,v) = \psi(x,y) = (x^2-y^2, x^2+y^2)$. Obsérvese que la imagen inversa bajo ψ de la recta $u = a > 0$ es una hipérbola y la imagen inversa bajo ψ de la recta $v = c > 0$ es un círculo. Demostrar que ψ no es inyectiva en $\mathbf{R}^2$, pero su restricción a $Q = \{(x,y): x > 0, y > 0\}$ es una aplicación inyectiva sobre $\{(u,v): v > |u|\}$. Sea φ el inverso de la restricción $\psi \mid Q$ demostrar que si $0 < a < b < c < d$, entonces φ aplica el rectángulo $A = [a,b] \times [c,d]$ en la región

$$\varphi(A) = \{(x,y): a \le x^2-y^2 < b, c < x^2+y^2 \le d\}.$$

Desmostrar que si $f: Q \to \mathbf{R}$ es continua, entonces

$$\iint_{\varphi(A)} f(x,y)\,d(x,y) = \iint_A f\left(\left(\frac{u+v}{2}\right)^{1/2}, \left(\frac{v-u}{2}\right)^{1/2}\right)\frac{1}{4(v^2-u^2)^{1/2}}\,d(u,v).$$

En particular, se tiene

$$\iint_{\varphi(A)} xy\,d(x,y) = \iint_A \tfrac{1}{8}\,d(u,v) = \tfrac{1}{8}(b-a)(d-c).$$

45.J. Sea $\psi: \mathbf{R}^2 \to \mathbf{R}$ como en el ejercicio anterior. Demostrar que ψ aplica la región triangular $\Delta = \{(x,y): 0 \le x \le 1, 0 \le y \le x\}$ en la región triangular

$$\Delta_1 = \psi(\Delta) = \{(u,v): 0 \le u \le 1, u \le v \le 2-u\}.$$

Aquí $J_\psi(x,y) = 8xy$. Si $\Omega_0 = (0,2) \times (0,2)$, y si f es continua en Δ_1, aplicar el teorema 45.11 para probar que

$$\iint_{\Delta_1} f(u,v)\,d(u,v) = \iint_\Delta f \circ \psi(x,y)\,|J_\psi(x,y)|\,d(x,y).$$

En particular, demostrar que

$$\iint_{\Delta}(x^2-y^2)(x^2+y^2)^{1/2}xy\,d(x,y)=\frac{1}{8}\iint_{\Delta_1}uv^{1/2}\,d(u,v).$$

45.K. Supóngase que $\alpha<\beta$ pertenecen a $[0,2\pi]$ y sea $h:[\alpha,\beta]\to \boldsymbol{R}$ continua y tal que $h(\theta)\ge 0$ para $\theta\in[\alpha,\beta]$. Sea $H=\{(\theta,r)\in\boldsymbol{R}^2:\alpha\le\theta\le\beta,\ 0\le r\le h(\theta)\}$ el conjunto ordenado de h (véase el ejercicio 44.0), de modo que H tiene contenido. La **curva polar** generada por h es la curva en $\boldsymbol{R}^2$ definida por $\theta\mapsto(h(\theta)\cos\theta,\ h(\theta)\operatorname{sen}\theta)$, y el **conjunto ordenado polar** de esta curva es el conjunto

$$H_1=\{(r\cos\theta,\ r\operatorname{sen}\theta)\in\boldsymbol{R}^2:\alpha\le\theta\le\beta,\ 0\le r\le h(\theta)\}.$$

Obsérvese que H_1 es la imagen de H bajo la aplicación polar (invertida) $\varphi_1(\theta,r)=(r\cos\theta,\ r\operatorname{sen}\theta)$ y usar el teorema 45.11 para probar que

$$c(H_1)=\frac{1}{2}\int_{\alpha}^{\beta}(f(\theta))^2\,d\theta.$$

45.L. Sea $a<b$ y supóngase que $f:[a,b]\to\boldsymbol{R}$ es continua y tal que $f(x)\ge 0$ para toda $x\in[a,b]$. Igual que en el ejercicio 44.O, sea $S_f=\{(x,y):a\le x\le b,\ 0\le y\le f(x)\}$ el **conjunto ordenado** de f. Defínase $\rho_x:\boldsymbol{R}^3\to\boldsymbol{R}^3$ como $\rho_x(x,y,\theta)=(x,\ y\cos\theta,\ y\operatorname{sen}\theta)$ y sea X_f la imagen de $S_f\times[0,2\pi]$ bajo ρ_x. (Al conjunto X_f se le llama el "sólido de revolución obtenido al girar el conjunto ordenado S_f en torno al eje x".) Usar el teorema 45.11 para probar que

$$c(X_f)=\pi\int_a^b(f(x))^2\,dx.$$

45.M. Sea $0\le a<b$ y sean $f:[a,b]\to\boldsymbol{R}$ y S_f como en el ejercicio anterior. Defínase $\rho_y:\boldsymbol{R}^3\to\boldsymbol{R}^3$ como $\rho_y(x,y,\theta)=(x\cos\theta,\ y,\ x\operatorname{sen}\theta)$ y sea Y_f la imagen de $S_f\times[0,2\pi]$ bajo ρ_y. (Al conjunto Y_f se le llama el "sólido de revolución obtenido al girar el conjunto ordenado S_f en torno al eje y") Usar el teorama 45.11 para probar que

$$c(Y_f)=2\pi\int_a^b xf(x)\,dx.$$

45.N. *(a)* Cambiando a coordenadas polares, demostrar que

$$\iint_{C_R}e^{-(x^2+y^2)}\,d(x,y)=\frac{\pi}{4}(1-e^{-R^2}),$$

en donde $C_R=\{(x,y):0\le x,\ 0\le y,\ x^2+y^2\le R^2\}$.

(b) Si $B_L=\{(x,y):0\le x\le L,\ 0\le y\le L\}$, demostrar que

$$\iint_{B_L}e^{-(x^2+y^2)}\,d(x,y)=\left(\int_0^L e^{-x^2}\,dx\right)^2.$$

(c) A partir del hecho de que $C_R\subseteq B_R\subseteq C_{R\sqrt{2}}$, demostrar que

$$\lim_{R\to\infty}\left(\int_0^R e^{-x^2}\,dx\right)^2=\frac{\pi}{4},$$

por lo que se infiere que $\int_0^{+\infty} e^{-x^2}\,dx = \frac{1}{2}\sqrt{\pi}$.

45.O. Sea $B=\{(x, y): 4x^2+9y^2 \leq 4\}$. Usar un cambio de variables apropiado para calcular

$$\iint_B e^{-(4x^2+9y^2)}\,d(x, y)=\frac{\pi}{6}(1-e^{-4}).$$

45.P. Observe que el conjunto $\{(x, y, z): 0 \leq x^2+y^2 \leq \frac{1}{2},\quad (x^2+y^2)^{1/2} \leq z \leq (1-x^2-y^2)^{1/2}\}$ un "corte de sector cónico de la bola unitaria" en $\boldsymbol{R}^3$. Obtener este conjunto como la imagen bajo la aplicación coordenada esférica Φ de la celda $[0, 1]\times[0, 2\pi]\times[0, \frac{1}{4}\pi]$.

(a) Demostrar que el contenido de este conjunto en $\boldsymbol{R}^3$ es igual a $\pi(2-\sqrt{2})/3$.

(b) Obtener el contenido de este conjunto usando la aplicación coordenada cilíndrica $\Gamma:(r, \theta, z) \to (x, y, z)=(r\cos\theta, r\operatorname{sen}\theta, z)$.

45.Q. Sea $a>0$ y sea A la intersección de los conjuntos

$$\{(x, y, z): x^2+y^2+z^2 \leq 4a^2\} \qquad \text{y} \qquad \{(x, y, z): z \geq a\}.$$

(a) Usar la aplicación coordenada esférica para probar que $c(A)=5\pi a^3/3$.

(b) Usar la aplicación coordenada cilíndrica para calcular $c(A)$.

45.R. Sea B la intersección de los conjuntos

$$\{(x, y, z): x^2+y^2+z^2 \leq 2\} \qquad \text{y} \qquad \{(x, y, z): x^2+y^2+z^2 \leq 2z\}.$$

(a) Usar la aplicación coordenada esférica para probar que $c(B)=\pi(4\sqrt{2}-3)/3$. $\boldsymbol{R}^p$.

(b) Usar la aplicación coordenada cilíndrica para calcular $c(B)$.

45.S. Sea $B_p(r)=\{x \in \boldsymbol{R}^p: \|x\| \leq r\}$ la bola con radio $r>0$ en el espacio $\boldsymbol{R}^p$. Se habrá de calcular el contenido $\omega_p(r)$ de $B_p(r)$.

(a) Usar un cambio de variables para probar que $\omega_p(r)=r^p\omega_p(1)$.

(b) Si $p \geq 3$, expresa la integral para $\omega_p(1)$ como una integral iterada y usar la parte *(a)* para demostrar que

$$\begin{aligned}\omega_p(1)&=\omega_{p-2}(1)\int_0^{2\pi}\left\{\int_0^1 (1-r^2)^{(p/2)-1} r\,dr\right\}d\theta\\&=\omega_{p-2}(1)2\pi/p.\end{aligned}$$

(c) Concluir que si $p=2k$ es par, entonces $\omega_p(1)=\pi^k/k!$ Si $p=2k-1$ es impar, entonces $\omega_p(1)=4^k\pi^{k-1}k!/(2k)!$ En términos de la función gama se tiene $\omega_p(1)=\pi^{p/2}/\Gamma(\frac{1}{2}p+1)$.

(d) Obtener el destacado resultado: $\lim(\omega_p(1))=0$.

45.T. Se habrá de obtener el resultado del ejercicio anterior de otra manera. Sea $p \in \boldsymbol{N}$ y supóngase que $\sigma: \boldsymbol{R}^p \to \boldsymbol{R}^p$ está definida como $\sigma(\theta)=\sigma(\theta_1, \ldots, \theta_p)=(\cos\theta_1, \theta_1\cos\theta_2, \ldots, \operatorname{sen}\theta_1\operatorname{sen}\theta_2\cdots\operatorname{sen}\theta_{p-1}\cos\theta_p)$.

(a) Demostrar que $\|\sigma(\theta)\|^2 \leq 1$ y que $\|\sigma(\theta)\|=1$ sólo cuando $\theta_j=0$ o $\theta_j=\pi$ para algún valor de $j=1, \ldots, p$.

(b) Demostrar que σ es un aplicación inyectiva de $(0, \pi)^p=(0, \pi)\times\cdots\times(0, \pi)$ (p veces) sobre el interior $\{x \in \boldsymbol{R}^p: \|x\|<1\}$ de la bola unitaria $B_p(1)$. Demostrar también que σ aplica $[0, \pi]^p$ sobre la bola unitaria, pero no es inyectiva en la frontera.

(c) Calculando el jacobiano de σ, se obtiene

$$J_\sigma(\theta) = (-1)^p(\operatorname{sen}\theta_1)^p(\operatorname{sen}\theta_2)^{p-2}\cdots(\operatorname{sen}\theta_{p-1})^2(\operatorname{sen}\theta_p).$$

Por lo tanto $J_\sigma(\theta) \neq 0$ para $\theta \in (0, \pi)^p$.

(d) Usando las fórmulas del producto de Wallis para $\int_0^{\pi/2} (\operatorname{sen}\theta)^k\, d\theta$ que se obtuvieron en el proyecto 30.γ, derivar las expresiones para $\omega_p(1)$ que se den en el ejercicio anterior

Proyecto

45.α. Este proyecto está basado en el proyecto **44.γ** y proporciona un acceso alternativo al teorema de cambio de variables 45.9. Sea $\Omega \subseteq \boldsymbol{R}^p$ abierto y supóngase que $\varphi: \Omega \to \boldsymbol{R}^p$ pertenece a la clase $C^1(\Omega)$, es inyectiva en Ω, y es tal que $J_\varphi(x) \neq 0$ para toda $x \in \Omega$. Para simplificar, también se supone que existe $M > 0$ tal que $\|\varphi(x) - \varphi(y)\| \leq M\,\|x - y\|$ para $x, y \in \Omega$.

(a) Si $\Phi: \mathscr{D}(\Omega) \to \boldsymbol{R}$ esta definida como

$$\Phi(A) = c(\varphi(A)) \qquad \text{para } A \in \mathscr{D}(\Omega),$$

entonces Φ es aditiva en $\mathscr{D}(\Omega)$ y tiene densidad fuerte igual a $|J_\varphi|$. Más aún, para alguna $M_1 > 0$, se tiene $\Phi(A) \leq M_1 c(A)$ para toda $A \in \mathscr{D}(\Omega)$.

(b) Si f es una función acotada, integrable en todo conjunto $\varphi(A)$, para $A \in \mathscr{D}(\Omega)$, y si $\Psi: \mathscr{D}(\Omega) \to \boldsymbol{R}$ esta definida como

$$\Psi(A) = \int_{\varphi(A)} f,$$

entonces Ψ es aditiva en $\mathscr{D}(\Omega)$. Más aún, para alguna $M_2 > 0$, se tiene $|\Psi(A)| \leq M_2 c(A)$ para toda $A \in \mathscr{D}(\Omega)$.

(c) Si f es una función acotada y continua en $\varphi(\Omega)$, y si Ψ está definida como en *(b)*, probar que Ψ tiene densidad fuerte igual a $(f \circ \varphi)\,|J_\varphi|$.

(d) Si f es como en *(c)*, demostrar que

$$\int_{\varphi(A)} f = \int_A (f \circ \varphi)\,|J_\varphi| \qquad \text{para } A \in \mathscr{D}(\Omega).$$

BIBLIOGRAFIA

Esta lista incluye libros y artículos que se citan en el libro y algunas referencias adicionales que serán útiles al profundizar los estudios.

Apostol, T.M., *Mathematical Analysis*, segunda edición, Addison Wesley, Reading, Mass., 1974.

Bartle, R.G., *The Elements of Integration*, Wiley, Nueva York, 1966.

Boas, R.P., Jr., *A primer of Real Functions*, Carus Monograph Number 13, Math. Assn. of America, 1960.

Bruckner, A.M., "*Differentiation of Integrals*", *Amer. Math. Monthly*, vol. 78, No 9, parte II, 1-51 (1971). (H. E. Slaught Memorial Paper, Number 12.)

Burkill, J. C. y H. Burkill, Mathematical *A Second Course in Analysis*, Cambridge Uni. Pres , Cambridge, 1970.

Cartan, H. P., *Cours de Mathematiques*, I. *Calcul Differentiel*; II. *Formes Differentielles*, Hermann, París, 1967. (traducción al ingles, Houghton-Mifflin, Boston, 1971.)

Cheney, E. L., *Introduction to Aproximation Theory*, McGraw-Hill, Nueva York, 1966.

Dieudonne, J. *Foundations of Modern Analysis*, Academic Press, Nueva York, 1960.

Dunford, N. y J. T. Schwartz, *Linear Operators*, parte I, Wiley-Interscience, Nueva York, 1958.

Finkbeiner, D. T., II, *Introduction to Matrices and Linear Transformations*, segunda edición, W. H. Freeman, San Francisco, 1966.

Gelbaum, B. R. y J. M. H. Olmsted, *Counterexamples in Analysis*, Holden-Day, San Francisco, 1964.

Halmos, P. R., *Naive Set Theory*, Van Nostrand, Princeton, 1960. (Republicado por Springer-Verlag, Nueva York, 1974.)

Hamilton, N. T. y J. Landin, *Set Theory*, Allyn-Bacon, Boston, 1961.

Hardy, G. H., J. E. Littlewood y G. Polya, *Inequalities*, segunda edición, Cambridge University Press, Cambridge, 1959.

Hewitt, E. y K. Stromberg, *Real and Abstract Analysis*, Springer- Verlag, Nueva York, 1965.

Hoffman, K. y R. Kunze, *Linear Albebra*, Prentice-Hall, Englewood Cliffs. 1961.

Kelley, J. L., *General Topology*, Van Nostrand, Nueva York, 1955.

Knopp, K., *Theory and Application of Infinite Series* (traducción al inglés) Hafner, Nueva York, 1951.

Lefschetz, S., *Introduction to Topology*, Princeton University Press, Princeton, 1949.

Luxemberg, W. A. J., "Arzela's Dominated Convergence Theorem for the Riemann Integral", *Amer. Math. Monthly*, Vol. 78, 970-979 (1971).

McShane, E. J., "A Theory of Limits", publicado en MAA *Studies in Mathematics*, Vol. 1, R. C. Buck, editor, Math. Assn. America, 1962.

"The Lagrange Multiplier Rule", *Amer. Math. Monthly*, Vol. 80, 922-925 (1973).

Royden, H. L., *Real Analysis*, segunda edición, Macmillan, Nueva York, 1968.

Rudin, W., *Principles of Mathematical Analysis*, segunda edición, McGraw-Hill, Nueva York, 1964.

Schwartz, J., "The Formula for Change of Variables in a Multiple Integral" *Amer. Math. Monthly*, Vol. 61, 81-85 (1954).

Simmons, G. F., *Introduction to Topology and Modern Analysis*, McGraw-Hill, Nueva York, 1963.

Spivak, M., *Calculus on Manifolds*, W. A. Benjamin, Nueva York, 1965.

Stone, N. H., "The Generalized Weierstrass Approximation Theorem", *Mathematic Magazine*, Vol. 21, 167-184, 237-254 (1947/48). (Reimpreso en MAA Studies in Mathematics, Vol. I, R. C. Buck, editor, Math. Assn, America, 1962.)

Suppes, P., *Axiomatic Set Theory*, Van Nostrand Princeton, 1961.

Titchmarsh, E. C., *The Theory of Functions*, segunda edición, Oxford University Press, Londres, 1939.

Varberg, D. E., "Change of Variables in Multiple Integrals", *Amer. Math. Monthly*, Vol. 78, 42-45 (1971).

Woll, J. W., Jr., *Functions of Several Variables*, Harcourt, Brace and World, Nueva York, 1966.

Wilder, R. L., *The Foundations of Mathematics*, Wiley, Nueva York, 1952.

SUGERENCIAS PARA EJERCICIOS SELECCIONADOS

Se recomienda al lector que no vea estas sugerencias a menos que encuentre dificultades. En muchos de los ejercicios se piden demostraciones y no hay una forma única que sea correcta; aun cuando el lector haya dado un argumento por completo distinto, éste puede ser absolutamente correcto. Sin embargo, con el objeto de ayudar al lector a aprender el material y a desarrollar sus habilidades técnicas, se ofrecen algunas sugerencias y unas cuantas soluciones. Se podrá ver que se dan más detalles para los primeros temas.

Sección 1

1.D. Por definición $A \cap B \subseteq A$. Si $A \subseteq B$, entonces $A \cap B \supseteq A$ de tal manera que $A \cap B = A$. Inversamente, si $A \cap B = A$, entonces $A \cap B \supseteq A$, por lo que se infiere que $B \supseteq A$.

1.E,F. La diferencia simétrica de A y B es la unión de $\{x : x \in A$ y $\{x : x \notin A\ x \notin B\}$ y $x \in B\}$.

1.H. Si x pertenece a $E \cap \bigcup A_j$, entonces $x \in E$ y $x \in \bigcup A_j$. Por lo tanto, $x \in E$ y $x \in A_j$ para al menos una j. Esto implica que $x \in E \cap A_j$ para al menos una j, de tal manera que

$$E \cap \bigcup A_j \subseteq \bigcup (E \cap A_j).$$

La inclusión opuesta se prueba invirtiendo estos pasos. La otra igualdad se maneja de manera análoga.

1.L. Si $x \in \mathscr{C}(\bigcap\{A_j : j \in J\})$, entonces $x \notin \bigcap\{A_j : j \in J\}$. Esto implica que existe $k \in J$ tal que $x \notin A_k$. Por lo tanto, $x \in \mathscr{C}(A_k)$, y entonces $x \in \bigcup\{\mathscr{C}(A_j) : j \in J\}$. Esto demuestra que $\mathscr{C}(\bigcap A_j) \subseteq \bigcup \mathscr{C}(A_j)$. La inclusión opuesta se demuestra invirtiendo estos pasos. La otra igualdad es análoga.

Sección 2

2.A. Si (a, c) y (a, c') pertenecen a $g \circ f$, entonces existen b, b' en B tales que (a, b), (a, b') pertenecen a f y (b, c), (b', c') pertenecen a g. Dado que f es una función, $b = b'$; y dado que g es una función, $c = c'$.

2.B. No, tanto $(0, 1)$ como $(0, -1)$ pertenecen a C.

2.D. Sea $f(x) = 2x$, $g(x) = 3x$.

2.E. Si (b, a), (b, a') pertenecen a f^{-1}, entonces (a, b), (a', b) pertenecen a f. Dado que f es inyectiva, $a = a'$. De donde f^{-1} es una función.

2.G. Si $f(x_1) = f(x_2)$ entonces $x_1 = g \circ f(x_1) = g \circ f(x_2) = x_2$. Por lo tanto, f es inyectiva.

2.H. Aplicar el ejercicio 2.G dos veces.

Sección 3

3.A. Sea $f(n) = n/2$, $n \in E$.

3.B. Sea $f(n) = (n+1)/2$, $n \in O$.

3.C. Sea $f(n) = n+1$, $n \in \mathbf{N}$.

3.E. Sea $A_n = \{n\}$, $n \in \mathbf{N}$. Entonces, cada conjunto A_n tiene sólo un punto, pero $\mathbf{N} = \bigcup\{A_n : n \in \mathbf{N}\}$ es infinito.

3.F. Si A es infinito y $B = \{b_n : n \in \mathbf{N}\}$ es un subconjunto de A, entonces la función definida como

$$\begin{aligned} f(x) &= b_{n+1}, \qquad x = b_n \in B, \\ &= x, \qquad x \in A \setminus B. \end{aligned}$$

es uno a uno y aplica A sobre $A \setminus \{b_1\}$.

3.H. Si f es una aplicación uno a uno de A sobre B y g es una aplicación uno a uno de B sobre C, entonces $g \circ f$ es una aplicación uno a uno de A sobre C.

Sección 4

4.G. Considere tres casos: $p = 3k$, $p = 3k+1$, $p = 3k+2$.

Sección 5

5.A. Dado que $a^2 \geq 0$ y $b^2 \geq 0, a^2 + b^2 = 0$ implica que $a^2 = b^2 = 0$.

5.D. Si $c = 1 + a$ con $a > 0$, entonces $c^n = (1+a)^n \geq 1 + na \geq 1 + a = c$.

5.G. Obsérvese que $1 < 2^1 = 2$. Si $k < 2^k$ para $k \geq 1$, entonces $k + 1 \leq 2k < 2 \cdot 2^k = 2^{k+1}$. Por lo tanto, $n < 2^n$ para toda $n \in \mathbf{N}$.

5.H. Obsérvese que $b^n - a^n = (b-a)(b^{n-1} + \cdots + a^{n-1}) = (b-a)p$, en donde $p > 0$.

5.M. $\{(x, y) : y = \pm x\}$.

5.N. Un cuadrado con vértices $(\pm 1, 0)$, $(0 \ \pm 1)$

Sección 6

6.A. Si $A = \{x_1\}$, entonces $x_1 = \sup A$. Si $A = \{x_1, \ldots, x_n, x_{n+1}\}$ y si $u = \sup\{x_1, \ldots, x_n\}$, demostrar que $\sup\{u, x_{n+1}\}$ es el supremo de A.

6.C. Sea $S = \{x \in \mathbf{Q} : x^2 < 2\}$.

6.E. De hecho, $\sup A \cup B = \sup\{\sup A, \sup B\}$.

6.H. Si $S = \sup\{f(x, y) : x \in X, y \in Y\}$, entonces $f(x, y) \leq S$ para toda $x \in X$, $y \in Y$, y entonces $f_1(x) \leq S$ para toda $x \in X$. De donde, $\sup\{f_1(x) : x \in X\} \leq S$. Inversamente, si $\varepsilon > 0$ existe (x_0, y_0) tal que $S - \varepsilon < f(x_0, y_0)$. De donde $S - \varepsilon < f_1(x_0)$ y por lo tanto $S - \varepsilon < \sup\{f_1(x) : x \in X\}$. Dado que $\varepsilon > 0$ es arbitraria, se deduce que $S \leq \sup\{f_1(x) : x \in X\}$.

6.K. Dado que $f(x) \leq \sup\{f(z) : z \in X\}$, se infiere que

$$f(x) + g(x) \leq \sup\{f(z) : z \in X\} + \sup\{g(z) : z \in X\}.$$

Por lo tanto, $\sup\{f(x)+g(x):x\in X\}$ es menor o igual al lado derecho. Análogamente, si $x\in X$, entonces

$$\inf\{f(z):z\in X\}+g(x)\le f(x)+g(x).$$

Si se usa 6.J, se deduce que

$$\inf\{f(z):z\in X\}+\sup\{g(x):x\in X\}\le \sup\{f(x)+g(x):x\in X\}.$$

Las otras afirmaciones se demuestran de manera análoga.

Sección 7

7.B. Sea $a\in A$; si $a\notin A'$, entonces $a\in B'$ y $\xi<\xi'\le$ una contradicción. Por lo tanto, $a\in A'$ y dado que $a\in A$ es arbitraria se tiene $A\subseteq A'$. Como $\xi<\xi'$, existe $x\in \boldsymbol{R}$ con $\xi<x<\xi'$. Como $\xi<x$, se debe tener $x\in B$. Pero como $x\in A'$, se infiere que $A\neq A'$.

7.C. Sean $A=\{x:x<1\}$, $B=\{x:x\ge 1\}$ y $A'=\{x:x\le 1\}$, $B'=\{x:x>1\}$.

7.E. Si $x\in I_n$ para toda n, se tiene una contradicción a la propiedad arquimediana 6.6.

7.F. Si $x\in J_n$ para toda n, se tiene una contradicción al corolario 6.7(b).

7.H. Todo elemento en F_1 tiene una expansión ternaria cuyo primer dígito es 0 ó 2. Los puntos en los cuatro subintervalos de F_2 tienen expansiones ternarias que empiezan

$$0.00\ldots,0.02\ldots,0.20\ldots,0.22\ldots,$$

y así sucesivamente.

7.J. Si n es suficientemente grande, $1/3^n<b-a$.

7.K. Tan cerca de 1 como se quiera.

Sección 8

8.E. La propiedad 8.3(ii) no se cumple.

8.H. El conjunto S_1 es el interior del cuadrado con vértices $(0,\pm1),(\pm1,0)$ y es el interior del cuadrado con vértices $(1,\pm1)$, $(-1,\pm1)$.

8.K. Tome $a=1/\sqrt{p}$, $b=1$.

8.L. Tome $a=1/p$, $b=1$.

8.M. Se tiene $|x\cdot y|\le\sum|x_i|\,|y_i|\le\{\sum|x_i|\}\sup|y_i|\le\|x\|_1\,\|y\|_1$ pero $|x\cdot y|\le p\,\|x\|_\infty\,\|y\|_\infty$ y si $x=y=(1,1,\ldots,1)$, se alcanza la igualdad.

8.N. La relación establecida implica que

$$\begin{aligned}\|x\|^2+2(x\cdot y)+\|y\|^2&=\|x+y\|^2=(\|x\|+\|y\|)^2\\&=\|x\|^2+2\,\|x\|\,\|y\|+\|y\|^2.\end{aligned}$$

De donde $x\cdot y=\|x\|\,\|y\|$ y la condición para la igualdad en el teorema 8.7 es válida siempre y cuando los vectores sean distintos de cero

8.P. Dado que $\|x+y\|^2=\|x\|^2+2(x\cdot y)+\|y\|^2$, la relación establecida se cumple si sólo si $x\cdot y=0$.

8.Q. Un conjunto K es convexo si y sólo si contiene al segmento de línea que une a cualesquiera dos puntos en K. Si $x,y\in K_1$, entonces $\|tx+(1-t)y\|\le t\,\|x\|+(1-t)\,\|y\|\le t+(1-t)=1$ de modo que $tx+(1-t)y\in K_1$ para

$0 \leq t \leq 1$. Los puntos $(\pm 1, 0)$ pertenecen e K_4, pero su punto medio (0,0) no pertenece a K_4.

8.R. Si x, y pertenecen a $\bigcap K_\alpha$, entonces, $x, y \in K_\alpha$ para toda α. De donde $tx + (1-t)y \in K_\alpha$ para toda α; de donde se deduce $\bigcap K_\alpha$ es convexo. Considere la unión de dos intervalos ajenos.

Sección 9

9.A. Si $x \in G$, sea $r = \inf\{x, 1-x\}$. Entonces, si $|y-x| < r$, se tiene $x - r < y < x + r$ por lo que $0 \leq x - r < y < x + r \leq 1$, y entonces $y \in G$. Si $z = 0$, entonces no existe un número real $r > 0$ tal que todo punto y en R que satisfaga y $|y| < r$ pertenece a F. Análogamente para $z = 1$.

9.B. Si $x \in G$, tomar $r = 1 - \|x\|$. Si $x \in H$, tomar $r = \inf\{\|x\|, 1 - \|x\|\}$. Si $z = (1, 0)$, entonces para cualquier $r > 0$, hay un punto y en $\mathscr{C}(F)$ tal que $\|y - z\| < r$.

9.G. Enumerar los puntos en el conjunto abierto con todos los números racionales coordenados. Después, proceder como en la demostración del teorema 9.11 usando bolas abiertas con centro en estos puntos racionales.

9.H. Argumentar como en el ejercicio anterior, pero esta vez usando bolas cerradas.

9.I. Tomar complementos y aplicar 9.H.

9.J. El conjunto A^0 es la unión de la colección de todos los conjuntos abiertos en A. Por lo tanto, cualquier conjunto abierto $G \subseteq A$ debe estar contenido en A^0. Por su definición, se debe tener $A^0 \subseteq A$. Se infiere que $(A^0)^0 \subseteq A^0$. Dado que A^0 es abierto y $(A^0)^0$ es la unión de todos los conjuntos abiertos en A^0, se debe tener $A^0 \subseteq (A^0)^0$. Por lo tanto, $(A^0)^0 = A^0$. Dado que $A^0 \subseteq A$ y $B^0 \subseteq B$ se infiere que $A^0 \cap B^0 \subseteq A \cap B$; pero como $A^0 \cap B^0$ es abierto, esto implica que $(A^0 \cap B^0) \subseteq (A \cap B)^0$. Por otro lado $(A \cap B)^0$ es un conjunto abierto y está contenido en A y en B; por lo tanto, $(A \cap B)^0 \subseteq A^0$ y $(A \cap B)^0 \subseteq B^0$, por lo que $(A \cap B)^0 \subseteq A^0 \cap B^0$. Por consecuencia, $A^0 \cap B^0 = (A \cap B)^0$. Dado que $\boldsymbol{R}^p$ es abierto, $(\boldsymbol{R}^p)^0 = \boldsymbol{R}^p$. Sea A el conjunto de todos los números racionales en (0, 1) y B el conjunto de todos los números irracionales en (0, 1). Entonces $A^0 \cup B^0 = \emptyset$, mientras que $(A \cup B)^0 = (0, 1)$.

9.L. Argumentar como en 9.J, o bien tomar complementos y usar 9.J.

9.N. Si $p = 1$, tomar $A = \boldsymbol{Q}$. En $\boldsymbol{R}^p$, tomar $\boldsymbol{Q}^p$.

9.0. Suponga que A, B son abiertos en $\boldsymbol{R}$. Sea $(x, y) \in A \times B$, de tal manera que $x \in A$ y $y \in B$. Existe $r > 0$ tal que si $|x' - x| < r$ entonces $x' \in A$ y $s > 0$ tal que si $|y' - y| < s$, entonces $y' \in B$. a ahora $t = \inf\{r, s\}$; la bola abierta con radio t está contenida en $A \times B$. El inverso es análogo.

Sección 10

10.C. Si x es un punto de acumulación de A en $\boldsymbol{R}^p$ y N es una vecindad x, entonces $N \cap \{y \in \boldsymbol{R}^p : \|y - x\| < 1\}$ contiene a un punto $a_1 \in A$, $a_1 \neq x$. El conjunto $N \cap \{y \in \boldsymbol{R}^p : \|y - x\| < \|a_1\|\}$ contiene a un punto $a_2 \in A$, $a_2 \neq x$ y además $a_2 \neq a_1$. Continuar este proceso.

10.F. Toda vecindad de x contiene una infinidad de puntos de $A \cup B$. Por lo que ya sea A o B (o posiblemente ambos) debe poseer un número infinito de elementos en esta vecindad.

Sección 11

11.A. Sea $G_n = \{(x, y): x^2 + y^2 < 1 - 1/n\}$ para $n \in \boldsymbol{N}$.

11.B. Sea $G_n = \{(x, y): x^2 + y^2 < n^2\}$ para $n \in \boldsymbol{N}$.

11.C. Sea $\mathcal{G} = \{G_\alpha\}$ una cubierta abierta para F y sea $G = \mathcal{C}(F)$, de tal manera que G sea abierta en $\boldsymbol{R}^p$. Si $\mathcal{G}_1 = \mathcal{G} \cup \{G\}$, entonces $\mathcal{G}_1$ es una cubierta abierta para K; por lo tanto, K tiene una subcubierta finita $\{G, G_\alpha, G_\beta, \ldots, G_\omega\}$. Entonces $\{G_\alpha, G_\beta, \ldots, G_\omega\}$ forma una subcubierta de $\mathcal{G}$ para el conjunto F.

11.D. Observe que si G es abierto en $\boldsymbol{R}$, entonces existe un subconjunto abierto G_1 de $\boldsymbol{R}^2$ tal que $G = G_1 \cap \boldsymbol{R}$. Alternativamente, usar el teorema de Heine-Borel.

11.E. Sea $\mathcal{G} = \{G_\alpha\}$ una cubierta abierta del intervalo unitario cerrado J en $\boldsymbol{R}^2$. Considere aquellos números reales x tales que el cuadrado $[0, x] \times [0, x]$ esté contenido en la unión de un número finito de conjuntos en $\mathcal{G}$ y sea x^* su supremo.

11.G. Sea $x_n \in F_n$, $n \in \boldsymbol{N}$. Si sólo hay un número finito de puntos en $\{x_n : n \in \boldsymbol{N}\}$, entonces al menos uno de ellos ocurre una infinidad de veces y es un punto común. Si hay una infinidad de puntos en el conjunto acotado $\{x_n\}$, entonces hay un punto de acumulación x. Dado que $x_m \in F_n$ para $m \geq n$ y dado que F_n es cerrado, entonces $x \in F_n$ para toda $n \in \boldsymbol{N}$.

11.H. Si $d(x, F) = 0$, entonces x es un punto de acumulación del conjunto cerrado F.

11.J. No. Sea $F = \{y \in \boldsymbol{R}^p : \|y - x\| = r\}$, entonces todo punto de F tiene la misma distancia a x.

11.K. Sea G un conjunto abierto y sea $x \in \boldsymbol{R}^p$. Si $H = \{y - x : y \in G\}$, entonces H es un conjunto abierto en $\boldsymbol{R}^p$.

11.M. Seguir el mismo argumento que en 11.7, sólo que usando celdas abiertas en vez de bolas abiertas.

11.Q. Suponga que $\boldsymbol{Q} = \bigcap \{G_n : n \in \boldsymbol{N}\}$, en donde G_n es abierto en $\boldsymbol{R}$. Por el teorema 6.10. El complemento F_n de G_n es un conjunto cerrado que no contiene a ningún subconjunto abierto no vacío. De donde, el conjunto de irracionales es la unión de una familia contable de conjuntos cerrados ninguno de los cuales contiene a un conjunto abierto no vacío, pero esto contradice al ejercicio 11.P.

Sección 12

12.B. Sea A, B una inconexión para $C' = C \cup \{x\}$. Entonces $A \cap C'$ y $B \cap C'$ son ajenos, no vacíos y tienen unión C'. Uno de estos conjuntos debe contener a x; suponga que es B. Dado que B es un conjunto abierto también contiene puntos de C de tal manera que $C \cap (B \setminus \{x\}) \neq \emptyset$. Pero entonces A, $B \setminus \{x\}$ forma una inconexión de C.

12.E. Modificar la demostración del teorema 12.4.

12.G. Por el teorema 12.8, los conjuntos C_1 y C_2 son intervalos. Fácilmente se puede ver que $C_1 \times C_2$ es convexo, de tal manera que 12.E es aplicable.

Sección 13

13.A. Analizar la posición geométrica de $iz = (-y, x)$ en términos de $z = (x, y)$.

13.B. Observe que $cz = (x \cos\theta - y \operatorname{sen}\theta, x \operatorname{sen}\theta + y \cos\theta)$, esto corresponde a una rotación en el sentido opuesto al de las manecillas del reloj de θ radianes en torno al origen.

13.C. El círculo $|z-c|=r$ se aplica en el círculo $|w-(ac+b)|=|a|\,r$. Se puede escribir $z=a^{-1}w-a^{-1}b$ y calcular $x=\text{Re}\,z$, $y=\text{Im}\,z$ en términos de $u=\text{Re}\,w$, $v=\text{Im}\,w$. Haciendo esto, con facilidad se puede ver que la ecuación $ax+by=c$ se transforma en una ecuación de la forma $Au+Bv=C$.

13.D. Un círculo queda fijado por g si y sólo si su centro está en el eje real. Las únicas rectas que quedan fijadas por g son los ejes real e imaginario.

13.E. Los círculos que pasan por el origen se mandan en rectas por h. Todas las rectas que no pasan por el origen se mandan en círculos que pasan por el origen, todas las rectas que pasan por el origen se mandan en rectas que pasan por el origen.

13.F. Todo punto de C, excepto el origen, es la imagen bajo g de dos elementos de C. Si $\text{Re}\,g(z)=k$, entonces $x^2-y^2=k$. Si $\text{Im}\,g(z)=k$, entonces $2xy=k$. Si $|g(z)|=k$, entonces $k\geq 0$ y $|z|=\sqrt{k}$.

Sección 14

14.B Observe que $0\leq \frac{1}{n}-\frac{1}{n+1}=\frac{1}{n(n+1)}\leq\frac{1}{n}$.

14.D Se tiene $0\leq \big|\|x_n\|-\|x\|\big|\leq\|x_n-x\|$.

14.I Sea $r\in\mathbf{R}$ tal que $\lim(x_{n+1}/x_n)<r<1$. Dado que el intervalo $(-1,r)$ es una vecindad de este límite, existe $K\in\mathbf{N}$ tal que $0<x_{n+1}/x_n<r$ para toda $n\geq K$. Demostrar ahora que $0<x_n<Cr^n$ para alguna C y $n\geq K$.

14.K. Considérense $(1/n)$ y (n).

14.L. Las sucesiones *(a)*, *(b)*, *(e)* y *(f)* convergen, las sucesiones *(c)* y *(d)* divergen.

14.M. Sea $r\in\mathbf{R}$ tal que $\lim(x_n^{1/n})<r<1$. Dado que el intervalo $(-1,r)$ es una vecindad de este límite, existe $K\in\mathbf{N}$ tal que $0<x_n^{1/n}<r$, por lo que $0<x_n<r^n$ para toda $n\geq K$.

Sección 15

15.A. Considere $z_n=y_n-x_n$ y aplicar el ejemplo 15.5 *(c)* y el teorema 15.6 *(a)*.

15.C. *(a)* Converge a 1. *(b)* Diverge. *(f)* Diverge.

15.D. Sea $\mathbf{Y}=-\mathbf{X}$.

15.F. Considere dos casos: $x=0$ y $x>0$.

15.G. Si.

15.H. Usar la sugerencia del ejercicio 15.F.

15.L. Observe que $b\leq x_n\leq b2^{1/n}$.

Sección 16

16.A. Por inducción, $1<x_n<2$ para $n\geq 2$. Dado que $x_{n+1}-x_n=(x_n-x_{n-1})/(x_nx_{n-1})$, la sucesión es monótona.

16.C. La sucesión es monótona y acotada. El límite es $(1+(1+4a)^{1/2})/2$.

16.D. La sucesión X es decreciente y acotada monótonamente

16.E. A un elemento x_k $X=(x_n)$ se le llama un "pico" para X si $x_k\geq x_n$ para $n>k$.

(i) Si hay una infinidad de picos con índices $k_1<k_2<\ldots$, entonces la sucesión (x_{k_j}) de picos es una subsucesión decreciente de $\mathbf{X}$.

(ii) Si sólo hay un número finito de picos con índices $k_1<\cdots<k_r$, sea $m_1>k_r$. Dado que x_{m_1} no es un pico, existen $m_2>m_1$ tal que $x_{m_1}<x_{m_2}$. Continuando de esta ma-

nera se obtiene una subsucesión estrictamente creciente (x_{m_j}) de X.

16.G. La sucesión es creciente y $x_n \leq n/(n+1) < 1$.

16.K. Existe $K \in \mathbf{N}$ tal que si $n \geq K$, entonces $L - \varepsilon \leq x_{n+1}/x_n \leq L + \varepsilon$. Usar ahora un argumento análogo al del ejercicio 14.I.

16.M. (a) e, (b) $e^{1/2}$, (c) *Sugerencia:* $(1+2/n) = (1+1/n)(1+1/(n+1))$, (d) e^3.

16.P. Sea $y_n \in F$ tal que $\|x - y_n\| < d + 1/n$. Si $y = \lim (y_{n_k})$, entonces $\|x - y\| = d$.

Sección 17

17.A. Todos.

17.C. Si $x \in \mathbf{Z}$, el límite es 1, si $x \notin \mathbf{Z}$, el límite es 0.

17.E. Si $x = 0$, el límite es 1, si $x \neq 0$, el límite es 0.

17.G. Si $x > 0$ y $0 < \varepsilon < \pi/2$, entonces $\tan(\pi/2 - \varepsilon) > 0$. Por lo tanto, $nx \geq \tan(\pi/2 - \varepsilon)$ para toda $n \geq n_x$, de donde $\pi/2 - \varepsilon \leq \operatorname{Arc\,tan} nx \leq \pi/2$.

17.H. Si $x > 0$, entonces $e^{-x} < 1$.

17.J. No necesariamente.

17.M. Considere la sucesión $(1/n)$ u observe que $\|f_n\|_D \geq \frac{1}{2}$.

17.P. Sí. 17.Q. Sí.

Sección 18

18.A. (a) ± 1. (b) 0. (c) ± 1. (d) ± 1.

18.E. Let m, $p \in \mathbf{N}$, $p \leq m$. Then $v_m(X+Y) = \sup\{x_n + y_n : n \geq m\} \leq \sup\{x_n : n \geq m\} + \sup\{y_n : n \geq m\} = v_m(X) + v_m(Y) \leq v_p(X) + v_m(Y)$. Por lo tanto,

$$(x+y)^* = \inf\{v_m(X+Y) : m \in \mathbf{N}\} \leq v_p(X) + y^*.$$

Dado que esto es válido para toda $p \in \mathbf{N}$, se infiere que $(x+y)^* \leq x^* + y^*$.

18.G. (a) $\pm\infty$. (c) 0, $+\infty$.

Sección 19

19.N. En

19.O. Aplicar el corolario 19.7 $x = \sup\{x_{mn} : m, n \in \mathbf{N}\}$.

19.P. Sea $x_{mn} = 0$ para $m < n$ y $x_{mn} = (-1)^m/n$ para $m \geq n$.

19.I. Si X es creciente y no converge en $\mathbf{R}$, entonces, X no es acotada.

19.K. *(a)* Ninguno existe. *(b, c)* Los tres son iguales. *(d)* Los límites iterados son

19.E. Si $j \leq n$, entonces $x_j \leq x_{n+1}$ y $x_j(1+1/n) \leq x_j + (1/n)x_{n+1}$. Ahora, sumar.

diferentes y el límite doble no existe. *(e)* El límite doble y un límite iterado son iguales. *(f)* Los límites iterados son iguales pero el límite doble no existe.

19.L. Sea $x_{mn} = n$ si $m = 1$ y $x_{mn} = 0$ if $m > 1$.

Sección 20

20.A. Si $a = 0$, tome $\delta(\varepsilon) = \varepsilon^2$. Si $a > 0$, usar la estima

$$|\sqrt{x} - \sqrt{a}| = \frac{|x-a|}{\sqrt{x}+\sqrt{a}} \leq \frac{|x-a|}{\sqrt{a}}.$$

20.B. Aplicar el ejemplo 20.5(*b*) y el teorema 20.6.

20.C. Aplicar el ejercicio 20.B y el teorema 20.6.

20.E. Demostrar que $|f(x) - f(\frac{1}{2})| = |x - \frac{1}{2}|$.

20.F. Todo número real es un límite de una sucesión de números racionales.

20.J. Existen sucesiones (x_n), (y_n) tales que $\lim(h(x_n)) = 1$, $\lim(h(y_n)) = -1$.

20.L. Demostrar que $f(a+h)-f(a) = f(h)-f(0)$. Si f es monótona en $\boldsymbol{R}$, entonces es continua en algún punto.

20.M. Demostrar que $f(0) = 0$ y $f(n) = nc$ para $n \in \boldsymbol{N}$. Además $f(n)+f(-n) = 0$, de tal manera que $f(n) = nc$ para $n \in \boldsymbol{Z}$. Dado que $f(m/n) = mf(1/n)$, al tomar $m = n$ se infiere que $f(1/n) = c/n$, por lo que $f(m/n) = c(m/n)$. Usar ahora la continuidad de f.

20.N. $g(0) = 0$, en cuyo caso $g(x) = 0$ para toda x en $\boldsymbol{R}$ o bien $g(0) = 1$, en cuyo caso $g(a+h)-g(a) = g(a)\{g(h)-g(0)\}$.

Sección 21

21.C. $f(1, 1) = (3, 1, -1)$, $f(1, 3) = (5, 1, -3)$.

21.D. Un vector (a,b,c) está en el rango de f si y sólo si $a-2b+c = 0$.

21.G. Si $\Delta = 0$, entonces $f(-b, a) = (0, 0)$. Si $\Delta \neq 0$, entonces la única solución de

$$ax + by = 0, \qquad cx + dy = 0$$

es $(x, y) = (0, 0)$.

21.I. Observe que $g(x) = g(y)$, si y sólo si $g(x-y) = \theta$.

21.P. Observe que $c_{ij} = e_i \cdot f(e_j)$ y aplicar la desigualdad de Schwarz.

Sección 22

22.C. Si $f(x_0) > 0$, entonces $V = \{y \in \boldsymbol{R} : y > 0\}$ en una vecindad de $f(x_0)$.

22.H. Sea $f(s, t) = 0$ si $st = 0$ y $f(s, t) = 1$ si $st \neq 0$.

22.L. Suponga que el coeficiente de la potencia más alta es positivo. Demostrar que existe $x_1 < 0 < x_2$ tal que $f(x_1) < 0 < f(x_2)$.

22.M. Sea $f(x) = x^n$. Si $c > 1$, entonces $f(0) = 0 < c < f(c)$.

22.N. Si $f(c) > 0$, hay una vecindad de c en la que f es positiva por lo que $c \neq \sup N$. Análogamente, si $f(c) < 0$.

22.O. Dado que f es estrictamente creciente y $a < b$, entonces f aplica al intervalo abierto (a, b) en una forma uno a uno sobre el intervalo abierto $(f(a), f(b))$, de donde se infiere que f^{-1} es continua.

22.P. Sí. Sean $a < b$ fijos y suponga que $f(a) < f(b)$. Si c es tal que $a < c < b$, entonces (i) $f(c) = f(a)$, o (ii) $f(c) < f(a)$, o (iii) $f(a) < f(c)$. El caso (i) se excluye por hipótesis. Si (ii), entonces existe a en (c, b) tal que $f(a_1) = f(a)$, una contradicción. Por lo tato, (iii) debe ser válido. De manera análoga, $f(c) < f(b)$ y f es estrictamente creciente.

22.Q. Suponga que g es continua y que $c_1 < c_2$ son los dos puntos en $\boldsymbol{I}$ en donde g alcanza un supremo. Si $0 < c_1$, elija números a_1, a_2 tales que $0 < a_1 < c_1 < a_2 < c_2$ y suponga que k satisface $g(a_1) < k < g(c_1)$. Entonces, existen tres números b_i tales que $a_1 < b_1 < c_1 < b_2 < a_2 < b_3 < c_2$ y en donde $k = g(b_i)$, una contradicción. Por lo tanto, se debe tener $c_1 = 0$ y $c_2 = 1$. Aplique ahora el mismo tipo de argumento a los puntos en donde g alcanza su ínfimo para obtener una contradicción.

22.S. Observe que $\varphi^{-1}(S)$ no es compacto. Además no es continua en (1,0).

Sección 23

23.A. Las funciones del ejemplo 20.5 *(a,b,i)* son uniformemente continuas en $\boldsymbol{R}$.

23.G. La función g es acotada y uniformemente continua en $[0, p]$.

23.I. Si (x_n) es una sucesión en $(0,1)$ con $x_n \to 0$, entonces $(f(x_n)$ es una sucesión de Cauchy y por lo tanto es convergente en $\boldsymbol{R}$.

23.K. Tome $f(x) = \text{sen } x$, $g(x) = x$, para $x \in \boldsymbol{R}$.

Sección 24

24.B. Tome $((1/n)f)$, en donde f es como en el ejemplo 20.5(g).

24.C. Obtener la función del ejemplo 20.5(h) de esta manera.

24.E. *(a)* La convergencia es uniforme en $[0, 1]$. *(b)* La convergencia es uniforme en cualquier conjunto cerrado que no contiene al 1. *(c)* La convergencia es uniforme en $[0, 1]$ o en $[c, +\infty)$, $c > 1$.

24.J. Se infiere que f es monótonamente creciente. Dado que f es uniformemente continua, si $\varepsilon > 0$, sea $0 = x_0 < x_1 < \cdots < x_h = 1$ tal que $f(x_j) - f(x_{j-1}) < \varepsilon$ y sea n_j tal que si $n \geq n_j$; entonces $|f(x_j) - f_n(x_j)| < \varepsilon$. Si $n \geq \sup\{n_0, n_1, \ldots, n_h\}$, demostrar que $|f(x) - f_n(x)| < 3\varepsilon$ para toda $x \in \boldsymbol{I}$.

24.S. Cualquier polinomio (o límite uniforme de una sucesión de polinomios) es acotado en un intervalo acotado.

Sección 25

25.G. *(b)* si $\varepsilon > 0$, existe $\delta(\varepsilon) > 0$ tal que si $c < x < c + \delta(\varepsilon)$, $x \in D(f)$, entonces $|f(x) - b| < \varepsilon$. *(c)* Si (x_n) es cualquier sucesión en $D(f)$ tal que $c < x_n$ y $c = \lim(x_n)$, entonces $b = \lim(f(x_n))$.

25.J. *(a)* Si $M > 0$, existe $m > 0$ tal que si $x \geq m$ y $x \in D(f)$, entonces $f(x) \gtrsim M$. *(b)* Si $M < 0$, existe $\delta > 0$ tal que si $0 < |x - c| < \delta$, entonces $f(x) < M$.

25.L. *(a)* Sea $\varphi(r) = \sup\{f(x) : x > r\}$ y fije $L = \lim_{r \to +\infty} \varphi(r)$. Alternativamente, si $\varepsilon > 0$, existe $m(\varepsilon)$ tal que si $x \geq m(\varepsilon)$, entonces $|\sup\{f(x) : x > r\} - L| < \varepsilon$.

25.M. Aplicar el lema 25.12

25.N. Considere la función $f(x) = -1/|x|$ para $x \neq 0$ y $f(0) = 0$.

25.P. Considere el ejemplo 20.5(h)

25.R. No necesariamente. Considere $f_n(x) = -x^n$ para $x \in \boldsymbol{I}$.

25.S. Sí

Sección 26

26.B. Demostrar que el conjunto $\mathscr{A}$ de polinomios en $\cos x$ satisface las hipótesis del teorema de Stone-Weierstrass.

26.E. Si $f(0) = f(\pi) = 0$, primero aproximar f por una función g que sea cero en algunos intervalos $[0, \delta]$ y $[\pi - \delta, \pi]$. Después, considérese $h(x) = g(x)/\text{sen } x$ para $x \in (0, \pi)$, $h(x) = 0$ para $x = 0, \pi$.

26.I. Considere $f(x) = \text{sen}(1/x)$ para $x \neq 0$.

26.K. Usar el teorema de Heine-Borel o el teorema de cobertura de Lebesgue como en la demostración del teorema de continuidad uniforme.

26.Q. *(a)* Dominio compacto, sucesión uniformemente equicontinua pero no acotada. *(b)* Dominio compacto, sucesión acotada pero no uniformemente equicontinua. *(c)* Dominio no compacto, sucesión acotada y uniformemente equicontinua.

Sección 27

27.D. Observe que $g'(0)=0$ y que $g'(x)=2x \operatorname{sen}(1/x)-\cos(1/x)$ para $x\neq 0$.

27.E. Sí.

27.L. Se puede escribir

$$\frac{f(x)-f(y)}{x-y}=\frac{x-c}{x-y}\cdot\frac{f(x)-f(c)}{x-c}-\frac{y-c}{x-y}\cdot\frac{f(y)-f(c)}{y-c}.$$

27.S. *(b)* Si $b\neq 0$, entonces, para $n\in N$ suficientemente grande, dada $x>n$, hay una $x_n>n$ tal que

$$|(f(x)-f(n))/x|=|(x-n)/x|\,|f'(x_n)|\geq|(x-n)/x|\,|b|/2.$$

Sección 28

28.F. Entre raíces consecutivas de p', el polinomio es estrictamente monotono. Si x_0 es una raíz de multiplicidad impar de p', entonces x_0 es un punto de extremo estricto para p.

28.H. La función f tiene raíces de multiplicidad n en $x=\pm 1$; f' tiene raices de multiplicidad $n-1$ en $x=\pm 1$, y una raíz simple dentro de $(-1, 1)$; etcétera.

28.O. Usar el ejercicio 27.O.

Sección 29

29.D. Si $\varepsilon>0$, entonces hay números racionales $r_1,\ldots,r_m$ en I tales que $0\leq f(x)<\varepsilon$ si $x\neq r_k$. Sea P una partición tal que cada uno de los subintervalos (a lo más $2m$) que contiene algunas de las $r_1,\ldots,r_m$ tiene longitud menor a $\varepsilon/2m$. Demostrar que $0\leq S(P;f,g)\leq 2\varepsilon$.

29.J. Si $f_1(x)=f(x)$ para $x\notin\{c_1,\ldots,c_m\}$ y $\varepsilon>0$, sea P una partición tal que cada uno de los subintervalos que contiene alguna de las $c_1,\ldots,c_m$ tiene longitud menor a $\varepsilon/2mM$, en donde $M\geq\sup\{\|f\|_J,\|f_1\|_J\}$. Usando los mismos puntos intermedios se tiene $|S(P;f,g)-S(P;f_1,g)|<\varepsilon$, en donde $g(x)=x$ para $x\in J$.

29.N. Suponga que $c\in(a,b)$; entonces f es g-integrable sobre $[a,c]$ y $[c,b]$. Si g_1 es la restricción de g a $[a,c]$, de 27.N se infiere que g_1 es continua en $[a,c]$; análogamente para la restricción g_2 de g a $[c,b]$. Del teorema 29.8 se infiere que fg_1' es integrable sobre $[a,c]$ y que fg_2' es integrable sobre $[c,b]$ y que

$$\int_a^c f\,dg=\int_a^c fg_1',\qquad \int_c^b f\,dg=\int_c^b fg_2'.$$

Sea ahora $(fg')(x)=f(x)g_1'(x)$ para $a\leq x\leq c$ y $(fg')(x)=f(x)g_2'(x)$ para $c<x\leq b$.

29.P. Si $\|P\|<\delta$ y si Q es un refinamiento de P, entonces $\|Q\|<\delta$.

29.R. Si $\varepsilon>0$, sea $P=(x_0,x_1,\ldots,x_n)$ una partición de J tal que si $P\supseteq P_\varepsilon$, y $S(P;f)$ es cualquier suma de Riemann correspondiente, entonces $|S(P;f)-\int_a^b f|<\varepsilon$.

Sea $M \geq \|f\|_J$ y sea $\delta = \varepsilon/4nM$. Si $Q = (y_0, y_1, \ldots, y_m)$ es una partición con norma $\|Q\| < \delta$, sea $Q^* = Q \cup P_\varepsilon$ de tal manera que $Q^* \supseteq P$ tiene a lo más $n-1$ puntos más que Q. Demostrar que $S(Q^*; f) - S(Q; f)$ se reduce a lo más a $2(n-1)$ términos de la forma $\pm\{f(\xi) - f(\eta)\}(x_j - y_k)$ con $|x_j - y_k| < \delta$.

Sección 30

30.C. Si $\varepsilon > 0$ está dada sea P_ε como en la demostración de 30.2. Si P es cualquier refinamiento de P_ε, entonces

$$|S(P_\varepsilon; f, g) - S(P; f, g)| \leq \sum |f(u_k) - f(v_k)|\,|g(x_k) - g(x_{k-1})|$$

en donde $|u_k - v_k| < \delta(\varepsilon)$ por lo que esta suma está dominada por εM. Usar ahora el criterio de Cauchy.

30.E. Un cálculo directo da

$$\left(\int_a^b (f(x))^n\, dx\right)^{1/n} \leq M(b-a)^{1/n}.$$

Inversamente, $f(x) \geq M - \varepsilon$ en algún intervalo de $[a, b]$.

30.H. Si $m \leq f(x) \leq M$ para $\alpha \leq x \leq \beta$, existe una A con $m \leq A \leq M$ tal que

$$F(\beta) - F(\alpha) = \int_\alpha^\beta f\, dg = A\{g(\beta) - g(\alpha)\}.$$

30.I. Sea $f(x) = -1$ para $x \in [-1, 0)$ y $f(x) = 1$ para $x \in [0, 1]$.

30.J. Aplicar el teorema del valor medio 27.6 para obtener $F(b)$-$F(a)$ como una suma de Riemann para la integral de f.

30.M. Si $m \leq f(x) \leq M$ para $x \in J$, entonces

$$m\int_a^b p \leq \int_a^b fp \leq M\int_a^b p.$$

Usar ahora el teorema de Bolzano 22.4.

30.P. Las funciones φ, φ^{-1} son uno a uno y continuas. Las particiones de $[c, d]$ están en correspondencia uno a uno con las particiones de $[a, b]$ y las sumas de Riemann-Stieltjes de $f \circ g$ con respecto a $g \circ \varphi$ están en correspondencia uno a uno con aquellas de f con respecto a g.

30.V.(a) $\frac{3}{4}$. (c) 9. (e) $\pi/2$.

Sección 31

31.K. Dado que $f_n(x) - f_n(c) = \int_c^x f_n'$, se puede aplicar el teorema 31.2 para obtener $f(x) - f(c) = \int_c^x g$ para toda $x \in J$. Demostrar que $g = f'$.

31.S. Aplicar el teorema 30.9 a (31.2) con $h(t) = (b-t)^{n-1}$.

31.V. Demostrar que las funciones G y H son continuas. El resto de la demostración es como en 31.9.

31.X. La función f es uniformemente continua en $J_1 \times J_2$.

31.Y. Sean $g_2(0) = 0$, $g_2(x) = \frac{1}{2}$ para $0 < x < 1$, y $g_2(x) = 1$.

Sección 32

32.D. *(a), (b), (d), (e)* son convergentes.

32.E. *(a)* es convergente para p, $q > -1$. *(b)* es convergente para $p+q > -1$.

32.F. *(a)* y *(c)* son absolutamente convergentes. *(b)* es divergente.

32.G. *(a)* es absolutamente convergente si $q > p+1$. *(b)* es convergente si $q > 0$ y absolutamente convergente si $q > 1$.

Sección 33

33.A. Si $0 \leq t \leq \beta$, entonces $x^t e^{-x} \leq x^\beta e^{-x}$.

33.B. Aplicar la prueba de Dirichlet 33.4.

33.C. *(a)* es uniformemente convergente para $|t| \geq a > 0$. *(b)* diverge si $t \leq 0$ y es uniformemente convergente si $t \geq c > 0$. (c) y (e) son uniformemente convergentes para toda t.

33.F. $\sqrt{\pi}$.

Sección 34

34.C. Agrupar los términos en la serie $\sum_{n=1}^{\infty} (-1)^n$ para producir convergencia a -1 y a 0.

34.G. Considerar $\Sigma\,((-1)^n n^{-1/2})$. Además, considerar el caso en que $a_n \geq 0$.

34.H. Si a, $b \geq 0$, entonces $2(ab)^{1/2} \leq a+b$.

34.I. Demostrar que $b_1+b_2+\cdots+b_n \geq a_1(1+\frac{1}{2}+\cdots+1/n)$.

34.J. Usar el ejercicio 34.F(a).

34.K. Demostrar que $a_1+a_2+\cdots+a_{2^n}$ está acotada por abajo por $\frac{1}{2}\{a_1+2a_2+\cdots+2^n a_{2^n}\}$ y por arriba por $a_1+2a_2+\cdots+2^{n-1}a_{2^{n-1}}+a_{2^n}$.

34.O. Considerar las sumas parciales s_k con $n/2 \leq k \leq n$ y aplicar el criterio de Cauchy.

Sección 35

35.C. *(a)* y *(e)* son divergentes. *(b)* es convergente.

35.D. *(b), (c)* y *(e)* son divergentes.

35.G. *(a)* es convergente. *(c)* es divergente.

35.L. Si $r < 1$, entonces $\log m < \log(m+1) - r/m$ para $m \in \mathbf{N}$. suficientemente grande. Demostrar que la sucesión $(x_n n \log n)$ es creciente.

Sección 36

36.A. Aplicar la prueba de Dirichlet.

36.D. *(a)* es convergente. *(b)* es divergente.

36.E. *(c)* Si $\Sigma\,(a_n)$ es absolutamente convergente, también lo es $\Sigma\,(b_n)$. Si $a_n = 0$ excepto cuando sen no está cerca de ± 1, se puede obtener un contra ejemplo. *(d)* Considerar $a_n = 1/n(\log n)^2$.

36.I. Si $m > n$, entonces $s_{mn} = +1$; si $m = n$, entonces $s_{mn} = 0$; si $m < n$, entonces $s_{mn} = -1$.

36.K. Observe que $2mn \leq m^2 + n^2$.

Sección 37

37.A. *(a)* y *(c)* convergen uniformemente para toda x. *(b)* converge para $x \neq 0$ y uniformemente para x en el complemento de cualquier vecindad de $x = 0$. *(d)* converge para $x > 1$ y uniformemente para $x \geq a$, en donde $a > 1$.

37.C. Si la serie es uniformemente convergente, entonces,

$$|c_n \sin nx + \cdots + c_{2n} \sin 2nx| < \varepsilon,$$

siempre que n sea suficientemente grande. Restringir ahora la atención a x en un intervalo tal que sen $kx > \frac{1}{2}$ para $n \leq k \leq 2n$.

37.H. (a) ∞, (c) $1/e$, (f) 1.

37.L. Aplicar el teorema de unicidad 37.17.

37.N. Demostrar que si $n \in \mathbf{N}$, entonces existe un polinomio P_n tal que si $x \neq 0$, entonces $f^{(n)}(x) = e^{-1/x^2} P_n(1/x)$.

37.T. Las series $A(x) = \Sigma\,(a_n x^n)$, $B(x) = \Sigma\,(b_n x^n)$, y $C(x) = \Sigma\,(c_n x^n)$ convergen a funciones continuas en I. Por el teorema de multiplicación 37.8, $C(x) = A(x)B(x)$ para $0 \leq x < 1$, y por continuidad $C(1) = A(1)B(1)$.

37.U. La sucesión de sumas parciales es creciente en el intervalo $[0, 1]$.

37.V. Si $\varepsilon > 0$, entonces $|a_n| \leq \varepsilon p_n$ para $n > N$. Descomponer la suma $\sum (a_n x^n)$ en una suma sobre $n = 1, \ldots, N$ y una suma sobre $n > N$.

Sección 38.

38.B. *(b)* Si $a_n = 0$, entonces

$$F(x + 2\pi) = \int_c^{x+2\pi} f(t)\,dt = \int_c^x f(t)\,dt + \int_x^{x+2\pi} f(t)\,dt$$
$$= F(x) + 0 = F(x)$$

para toda $x \in \mathbf{R}$.

38.E. *(c)* Calcular $f_2(0)$ de dos maneras.

38.G. *(a)* Si k_1 fuera continua, entonces $k_1(-\pi) = -\pi^3$ pero dado que k_1 tiene período 2π, $k_1(-\pi) = k_1(\pi) = \pi^3$.

38.I. (b) $\dfrac{2}{\pi} - \dfrac{4}{\pi}\left[\dfrac{\cos 2x}{1 \cdot 3} + \dfrac{\cos 4x}{3 \cdot 5} + \dfrac{\cos 6x}{5 \cdot 7} + \cdots\right]$.

(c) $\dfrac{1}{2} + \dfrac{2}{\pi}\left[\dfrac{\cos x}{1} - \dfrac{\cos 3x}{3} + \dfrac{\cos 5x}{5} - \cdots\right]$.

(e) $\dfrac{\pi^2}{6} - \left[\dfrac{\cos 2x}{1^2} + \dfrac{\cos 4x}{2^2} + \dfrac{\cos 6x}{3^2} + \cdots\right]$.

38.K. (b) $\dfrac{8}{\pi}\left[\dfrac{\text{sen}\, 2x}{1 \cdot 3} + \dfrac{2\,\text{sen}\, 4x}{3 \cdot 5} + \dfrac{3\,\text{sen}\, 6x}{5 \cdot 7} + \cdots\right]$.

(e) $\dfrac{8}{\pi}\left[\dfrac{\text{sen}\, x}{1^3} + \dfrac{\text{sen}\, 3x}{3^3} + \dfrac{\text{sen}\, 5x}{5^3} + \cdots\right]$.

38.N. *(d)* Usar el ejercicio 38.G(b).

38.R. (a) $\frac{4}{\pi}\left[\frac{\operatorname{sen}\frac{1}{2}\pi x}{1}-\frac{\operatorname{sen}\pi x}{2}+\frac{\operatorname{sen}\frac{3}{2}\pi x}{3}-\cdots\right]$.

(b) $1+\frac{4}{\pi}\sum_{n=1}^{\infty}\left[-\frac{1+(-1)^{n+1}}{n^2\pi}\cos\frac{n\pi x}{4}+\frac{(-1)^{n+1}}{n}\operatorname{sen}\frac{n\pi x}{4}\right]$.

38.S. Usar el teorema de Fejer 38.12 y el teorema 19.3. 38.T. Modificar las demostraciones de 38.7 y 38.12.

Sección 39

39.G. Se tiene $|G(u, v)-G(0,0)| \le |u^2+v^2| = \|(u, v)\|^2$ tal que $DG(0,0)(u, v)=0$. si $(x, y) \ne (0,0)$, entonces $D_1G(x, x)=2x \sin(2x^2)^{-1}-x^{-1}\cos(2x^2)^{-1}$, que no es acotada $x \to 0$. cuando

39.L. (a) $\nabla_{(a,b,c)}f_1=(2a, 2b, 2c)$. (c) $\nabla_{(a,b,c)}f_3=(bc, ac, ab)$.

39.M. *(a)* 0. *(c)* $4/\sqrt{6}$.

39.O. *(a)* en (1,2) se tiene $\{(x, y, z): z-5=2(x-1)+4(y-2)\}$. *(c)* en (1,1) se tiene $\{(x, y, z): z-\sqrt{2}=-(x+y-2)/\sqrt{2}\}$.

39.Q. *(a)* En $t=0$ se tiene $\{(x, y, z): x=t, y=0, z=0\}$; en $t=1$ se tiene $\{(x, y, z): x=1+s, y=1+2s, z=1+3s\}$. *(c)* en $t=\frac{1}{2}\pi$ se tiene $\{(x, y, z): x=-2s,\ y=2,\ z=\frac{1}{2}\pi+s\}$.

39.S. *(b)* En el punto $(3,-1,-3)$ correspondiente a $(s, t)=(1,2)$ se tiene $S_h=\{(x, y, z): x=3+(s-1)+(t-2),\ y=-1+(s-1)-(t-2),$

(d) En el punto (1,0,0) correspondiente a $(s, t)=(0,\frac{1}{2}\pi)$ se tiene $S_h=\{(x, y, z): x=1, y=s, z=-(t-\frac{1}{2}\pi)\}$.

39.V. Obsérvese que si $y \in \mathbf{R}^q$, $z \in \mathbf{R}^r$, entonces $(y, z) \in \mathbf{R}^q \times \mathbf{R}^r = \mathbf{R}^{q+r}$ es tal que $\|(y, z)\|^2=\|y\|^2+\|z\|^2$.

Sección 40

40.A. $F'(t)=2(3t+1)3+2(2t-3)2=26t-6$.

40.D. $D_1F(s, t)=(\operatorname{sen} s \cos t+\operatorname{sen} t)(-\operatorname{sen} s)+(\cos s+\operatorname{sen} t)(\cos s \cos t)+0$.

40.G. (a) $D_1F(x, y)=f'(xy)y$, $D_2F(x, y)=f'(xy)x$.

(d) $D_1F(x, y)=f'(x^2-y^2)(2x)$, $D_2F(x, y)=f'(x^2-y^2)(-2y)$.

40.K. *(b)* Dado que $g'(t)=D_1f(tc)c_1+\cdots+D_pf(tc)c_p$, de la relación de Euler se infiere que

$$tg'(t)=(tc_1)D_1f(tc)+\cdots+(tc_p)D_pf(tc)=kf(tc)=kg(t).$$

Por lo tanto (¿por qué?) $g(t)=Ct^k$ para alguna constante C. Dado que $f(c)=g(1)=C$, se deduce que $f(tc)=g(t)=t^kf(c)$, por lo que f es homogénea de grado k.

40.M. Dado que

$$\|B(x+u, y+v)-B(x, y)-(B(x, v)+B(u, y))\| = \|B(u, v)\| \le M\,\|u\|\,\|v\| \le \tfrac{1}{2}M(\|u\|^2+\|v\|^2)=\tfrac{1}{2}M\,\|(u, v)\|^2,$$

se infiere que $DB(x, y)(u, v)$ existe y es igual a $B(x, v)+B(u, y)$.

40.P. Dado que $Dg(c)(u)=(ug_1'(c), \ldots, ug_p'(c))=ug'(c)$ para $u \in \mathbf{R}$, de la regla de la cadena se infiere

$$Dh(c)(u) = Df(g(c))(Dg(c)(u)) = Df(g(c))(ug'(c)) = uDf(g(c))(g'(c))$$

por lo que $h'(c) = Df(g(c))(g'(c))$.

40.Q. Si $f = (f_1, \ldots, f_q)$, existen puntos $c_i \in S$ tales que $f_i(b) - f_i(a) = Df_i(c_i)(b-a)$. Supóngase ahora que L tiene la representación matricial $[D_j f_i(c_i)]$.

40.R. Por el teorema 12.7 cualesquiera dos puntos en Ω se pueden unir por una curva poligonal que esté dentro de Ω. Aplicar el teorema del valor medio a cada segmento de esta curva.

40.U. De hecho $D_x f(x, y) = y(x^2 - y^2)(x^2 + y^2)^{-1} + 4x^2 y^3 (x^2 + y^2)^{-2}$ y $D_{yx} f(0, 0) = -1$, mientras que $D_{xy} f(0, 0) = +1$.

40.W. Si $\varphi : (-\varepsilon, 1 + \varepsilon) \to \boldsymbol{R}^q$ se define como $\varphi(t) = f(a + t(b-a))$, entonces $\varphi'(t) = Df(a + t(b-a))(b-a)$. Escribir $\varphi(t) = (\varphi_1(t), \ldots, \varphi_q(t))$ en donde $\varphi_j(t) = f_j(a + t(b-a))$ y obsérvese que $\varphi_j(1) - \varphi_j(0) = \int_0^1 \varphi_j'(t)$.

Sección 41

41.A. Usar el ejercicio 21.P.

41.D. Aquí $Df(0) = 0$. No.

41.E. Considerar los ejercicios 27.H y 22.O.

41.F. Considerar el ejercicio 40.L.

41.J. Encontrar los extremos relativos cerca de 0.

41.Q. *(a)* En (1,1,2) se tiene $S_F\{(x, y, z) : 2x + 2y - z = 2\}$. *(c)* En $(4, \frac{1}{2}, 2)$ se tiene $S_F = \{(x, y, z) : x + 8y - 2z = 4\}$.

41.S. Su $D_1 f$ es cero en un conjunto abierto, aplicar el ejercicio 40.S. Si $D_1 f(x_0, y_0) \neq 0$, considerar $F(x, y) = (f(x, y), y)$ cerca de (x_0, y_0).

41.T. Si $D_1 g(c) \neq 0$, considerar $G(x, y) = g(x) + (0, y)$.

41.V. Demostrar, igual que en la demostración de 41.6, que si $\|y\| < m/2$, entonces hay un vector $x \in \boldsymbol{R}^p$ con $\|x\| \leq 1$ tal que $y = L_1(x)$.

41.W. Si $y \in \boldsymbol{R}^p$, sea $x_0 = 0$

$$x_{n+1} = x_n - (f(x_n) - f(x_{n-1}) - (x_n - x_{n-1})).$$

demostrar, igual que en la demostración de 41.6, que $\bar{x} = \lim (x_n)$ existe y $f(\bar{x}) = y$.

Sección 42

42.A. *(a)* Punto silla en (0,0). *(b)* Mínimo relativo estricto en $(-2, \frac{1}{2})$. *(c)* Punto silla en (0, −1); mínimo relativo estricto en (0,3). *(f)* Punto silla en (0,0); mínimos relativos estrictos en (0, −1) y (0,2).

42.D. Si f no es constante, entonces ya sea el supremo o el ínfimo de f en $S = \{x \in \boldsymbol{R}^p : \|x\| \leq 1\}$ es distinto de 0. Dado que S es compacto, este supremo (o ínfimo) se alcanza en un punto $c \in S$. La hipótesis descarta la posibilidad de que $\|c\| = 1$.

42.F. *(a, d)* Mínimo relativo en (0,0). *(b, c, e)* Punto silla en (0,0). *(f)* Mínimo relativo estricto en (0,0).

42.G. Punto silla en (1,1).

42.H. Los monos tienen cola.

42.I. $\frac{7}{3}$.

42.K. $\frac{2}{9}$.

42.S. *(a)* Valor máximo $=1$, alcanzado en $(\pm 1, 0)$; valor mínimo $=-1$, alcanzado en $(0, \pm 1)$. *(b)* Valor máximo $=3$, alcanzado en $(1,0)$; valor mínimo $=-1$, alcanzado en $(-1, 0)$. *(c)* Valor máximo $=4$, alcanzado en $(1, \pm 1)$; valor mínimo $=-1$, alcanzado en $(-1, 0)$. *(d)* Valor máximo $=1$, alcanzado en $(0, \pi/2)$; valor mínimo $=-1$; al alcanzado en $(0, -\pi/2)$.

42.U. Valor máximo $=1$, alcanzado en $(1,0,0)$; valor mínimo $=\frac{5}{9}$, alcanzado en $(-\frac{1}{3}, \frac{2}{3}, \frac{2}{3})$.

Sección 43

43.B. Si $p \in \boldsymbol{N}$ está dada sea $n > (2^{1/p}-1)^{-1}$. Si una celda I en $\boldsymbol{R}^p$ tiene longitudes laterales $0 < a_1 \leq a_2 \leq \cdots \leq a_p$, sea $c = a_1/n$. Entonces, I está contenida en la unión de $n([a_2/c]+1)\cdots([a_p/c]+1)$ cubos con longitud lateral c que tienen contenido total menor a $2(a_1 \cdots a_p) = 2c(I)$. De donde, si Z está contenido en la unión de celdas con contenido total menor a ε, está contenido en la unión de cubos con contenido total menor a 2ε.

43.E. No.

43.H. Si la cerradura de J_j es $[a_{j1}, b_{j1}]\times\cdots\times[a_{jp}, b_{jp}]$, para $j=1, \ldots, n$, y si $I=[a_1, b_1]\times\cdots\times[a_p, b_p]$, sea P_1 la partición de $[a_1, b_1]$ que se obtiene al usar los puntos $\{a_{j1}, b_{j1} : j=1, \ldots, n\}, \ldots,$ y P_p la partición de $[a_p, b_p]$ que se obtiene al usar $\{a_{jp}, b_{jp} : j=1, \ldots, n\}$. Las particiones $P_1, \ldots, P_p$ inducen una partición de I.

43.I. Encerrar a Z en la unión de un número finito de celdas cerradas en I con contenido total menor a ε. Aplicar ahora 43.H.

43.J. Encerrar a Z en la unión de un número finito de celdas abiertas en I con contenido total menor a ε. Aplicar ahora 43.H.

43.K. Se forma una sucesión de particiones de I en cubos con longitud lateral $2^{-n}\delta$ por medio de una bisección sucesiva de los lados de I. Dado un cubo $K \subseteq I$ con longitud lateral r, encerrar a K en la unión de todos los cubos en la n-ésima partición que tengan intersección no vacía con K. Si n es tan grande que, $(1+\delta/2^{n-1}r)^p < 2$, entonces esta unión tiene contenido total menor a $2c(K)$.

43.L. Usar 43.G.

43.M. Usar 43.L.

43.R. Primero tratar el caso $f=g$; después considerar $(f+g)^2$.

43.T. Sea $M > \|f\|_I, \|g\|_I$. Dado que f y g son uniformemente continuas en K, si P_ε es suficientemente fina, entonces f y g varían menos de $\varepsilon/2M$ en cada K_j tal que para cualquier $p_j \in K_j$ se tiene $|\int_K fg - \sum f(p_j)g(p_j)c(K_j)| \leq (\varepsilon/2)c(K)$. Entonces se tiene

$$\left|\int_K fg - \sum f(x_j)g(y_j)c(K_j)\right| \leq \left|\int_K fg - \sum f(x_j)g(x_j)c(K_j)\right| + \left|\sum f(x_j)[g(x_j)-g(y_j)]c(K_j)\right| \leq \varepsilon c(K).$$

43.V. *(d)* Si Z es compacto y está contenido en la unión de celdas abiertas $J_1, J_2, \ldots$, entonces está contenido en la unión de un número finito de estas celdas.

Sección 44

44.B. *(a)* Si $c \notin b(A)$, entonces c es un punto interior de A o bien es un punto inte-

rior de $\mathscr{C}(A)$. En cualquiera de los casos, hay una vecindad de c, ajena a $b(A)$, de tal manera que $\mathscr{C}(b(A))$ es abierto.

44.C. En el ejemplo 43.2(g), se tiene $S^- = b(S) = I \times I$. Sin embargo, $I \times \{0, 1\} \cup \{0, 1\} \times I$.

44.D. Dado que $(A \cap B)^- \subseteq A^- \cap B^-$, se infiere que

$$\begin{aligned} b(A \cap B) &= (A \cap B)^- \cap (\mathscr{C}(A \cap B))^- \subseteq A^- \cap B^- \cap (\mathscr{C}(A) \cup \mathscr{C}(B))^- \\ &= A^- \cap B^- \cap (\mathscr{C}(A)^- \cup (\mathscr{C}(B)\ ) \\ &= (B^- \cap b(A)) \cup (A^- \cap b(B)) \subseteq b(A) \cup b(B). \end{aligned}$$

44.J. Si $\varepsilon > 0$, sea P_ε una partición como en el ejercicio 43.P y tal que la unión de las celdas en P_ε que contienen puntos en $b(A)$ tiene contenido total menor a $\varepsilon/2\,\|f\|_I$. Aplicar Ahora el ejercicio 43.P a la restricción de f a A.

44.K. Dado que $mg(x) \leq f(x)g(x) \leq Mg(x)$ para $x \in A$, se infiere que $m\int_A g \leq \int_A$ $M \int_A g$. Si $\int_A g \neq 0$, tomar $\mu = (\int_A fg)(\int_A g)^{-1}$.

44.N. El conjunto $(K) = K$ no tiene contenido cero.

44.R. Observe que $F(x, y) = \int_0^y \{\int_0^x f(s, t)\, ds\}\, dt$.

Sección 45

45.A. Analizar las demostraciones de 45.1-45.4.

45.D. *(a)* 6π.

45.F. $(e-1)^2$.

45.H. Sea $u = xy$, $v = y/x^2$. El área es igual a $(\log 2)/3$.

Indice

i516 (1469x2409x2 pcx)

R

S

—o◯o—